DOCUMENT RESUME

ED 434 637 HE 032 415

TITLE	Assessment Update: The First Ten Years.
INSTITUTION	National Center for Higher Education Management Systems, Boulder, CO.
PUB DATE	1999-00-00
NOTE	530p.; "Featuring Complete Collections of Columns from: Trudy W. Banta, Editor, 'Editor's Notes'; Peter T. Ewell, 'From the States'; Jeffrey A. Seybert, 'Community College Strategies'; Peter J. Gray, 'Campus Profiles'; Gary R. Pike, 'Assessment Measures'."
PUB TYPE	Collected Works - General (020)
EDRS PRICE	MF02/PC22 Plus Postage.
DESCRIPTORS	*College Outcomes Assessment; Community Colleges; *Educational Assessment; Educational Policy; *Educational Trends; Evaluation Methods; Higher Education; Opinions; *Program Evaluation; Trend Analysis

ABSTRACT

This document provides complete collections of columns from the journal "Assessment Update" for the past 10 years. By-lined columns from Trudy W. Banta, Editor, include articles on: (1) encouraging faculty involvement in assessment; (2) the growing variety of assessment methods; (3) assessment in response to external pressures: performance funding in Tennessee; (4) some national assessment issues; and (5) international perspectives on assessing quality in education. Thirty-one articles from Peter T. Ewell include: (1) "The Current Pattern of State Assessment Initiatives"; (2) "Mandated Assessment: A Look in the Cultural Mirror"; (3) "'Program Excellence' in Ohio: An Indirect Approach to Assessment"; (4) "So Are They Really Going Away"; and (5) "Technology Facilitates National Data Collection Strategies." Fifteen articles from Jeffrey A. Seybert include: (1) "How To Initiate an Assessment Program"; (2) "Assessment of Noncredit Continuing Education and Community Service Programs and Courses"; (5) "Assessment at Two-Year Institutions"; and (4) "A New Paradigm for Evaluating Transfer Success." The 33 "Campus Profiles", written by Peter J. Gray, present case studies that suggest that in order for assessment to be successfully institutionalized on a campus, they must be made a priority over an extended period by central and departmental leaders. Thirty-two articles by Gary R. Pike include articles on general education, critical thinking, surveys, writing assessment, major field, assessment resources, and enrollment management. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 434 637

ASSESSMENT UPDATE

THE FIRST TEN YEARS

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

C. Roberts

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

TRUDY W. BANTA, EDITOR

HE 032 415

BEST COPY AVAILABLE

ASSESSMENT UPDATE:
The First Ten Years

Featuring Complete Collections of Columns from

Trudy W. Banta, Editor, *Editor's Notes*
Peter T. Ewell, *From the States*
Jeffrey A. Seybert, *Community College Strategies*
Peter J. Gray, *Campus Profile*
Gary R. Pike, *Assessment Measures*

National Center for Higher Education Management Systems (NCHEMS)
P.O. Box 9752 Boulder, Colorado 80301-9752

All articles originally appeared in *Assessment Update*.

Copyright © 1999 Jossey-Bass Inc., Publishers

No part of this publication may be reproduced, stored in a
retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior written permission of
Jossey-Bass Inc., Publishers, 350 Sansome Street, San Francisco, CA 94104.

Assessment Update: The First Ten Years

Preface

In 1998 Gale Erlandson, higher education editor at Jossey-Bass Publishers, asked if I would be interested in editing a newsletter on assessment. At that time Tennessee, Virginia, New Jersey and Colorado had attracted the attention of many other states with their assessment initiatives aimed at public colleges and universities and the Southern Association of Colleges and Schools was blazing a path for other regional accrediting associations in its implementation of an institutional effectiveness criterion. A couple of *New Directions* issues on assessment were fairly flying off the shelves at Jossey-Bass, so Gale Erlandson had surveyed the marketplace and decided that this fast-developing new field needed a newsletter.

The idea of assembling a staff to produce a periodical that would appear several times a year was daunting. But I received strong support for this effort from the administration at the University of Tennessee, and the same was true when I moved in 1992 to Indiana University-Purdue University Indianapolis. So in January 1989 the first issue of *Assessment Update* was published, and we have never missed a Jossey-Bass deadline. We would not have been able to establish, and then sustain, the quality of *Assessment Update* over the first decade of its history without the contributions of columns by a committed group of experienced professionals in the field of assessment. When first approached, Peter Ewell, Peter Gray, and Gary Pike, and later Jeff Seybert, generously agreed to contribute regular columns. A few of the early issues would have had to go to press with empty pages had we not had the columns to add! Over the years readers have learned to turn regularly to Peter Ewell's *From the States* for the latest developments in assessment in the states and at the national level, to Peter Gray's *Campus Profiles* for institutional case studies, to Gary Pike's *Assessment Measures* for helpful information about instruments and methods, and to Jeff Seybert's *Community College Strategies* for developments in the two-year sector.

Marcia Mentkowski, Pat Hutchings, Tom Angelo, John Harris, and Kay McClenney have made their own generous contributions to *Assessment Update* over the years. As they have met with the columnists in annual meetings of the Editorial Board, I have often heard one of them make this kind of remark, "Even though it is a newsletter, the columns and articles in *Assessment Update* contain information that does not go out of date. You really ought to put together some collections of related materials over time." When *Assessment Update* was eight, conversation in the board meeting provided the impetus for producing a volume containing all the columns, including the *Editor's Notes*. In the ninth year my colleagues and I at IUPUI simply copied the columns, added introductions written by the columnists, bound the volume, and offered 25 copies for sale at the 1997 Assessment Conference in Indianapolis. All the copies were purchased within 20 minutes!

The success of our trial balloon encouraged us to find a publisher for a collection of the columns from a full decade of *Assessment Update* issues. Peter Ewell said NCHEMS would be interested. Gale Erlandson expedited the permission-granting process at Jossey-Bass and Clara Roberts at NCHEMS stepped forward to assume responsibility for what turned out to be a fairly arduous process of producing the collection in its present form.

To Peter, Gary, Peter, and Jeff, thank you for your friendship and your unfailing support during *Assessment Update*'s first decade—and beyond. To Gale and Clara, thank you for helping us realize our ambition to make this volume available to our colleagues around the world who are engaged in their own work of improving teaching, learning, and campus services through assessment. My fellow columnists and I are grateful for this opportunity to reflect on developments in assessment over the past decade and sincerely hope that our readers will find helpful this history of the development of the field and related issues at the classroom, departmental, campus, state, regional, national, and international levels.

Trudy W. Banta
Indianapolis
April 1999

Contents

Page

Preface vii

Acknowledgements ix

Editor's Notes
Trudy W. Banta

About the Author 2
Contents 3
Reflections on Ten Years as Editor of *Assessment Update* 5

Section I
Encouraging Faculty Involvement in Assessment 7

Section II
The Growing Variety of Assessment Methods 29

Section III
Assessment in Response to External Pressures: Performance Funding in Tennessee 93

Section IV
Some National Assessment Issues 103

Section V
International Perspectives on Assessing Quality in Higher Education 129

From the States
Peter T. Ewell

About the Author 148
Contents 149
From the States: Ten Turbulent Years 151

Page

Community College Strategies
Jeffrey A. Seybert

About the Author 242
Contents 243
Community College Strategies 245

Campus Profiles
Peter J. Gray

About the Author 294
Contents 295
Campus Profiles 297

Assessment Measures
Gary R. Pike

About the Author 418
Contents 419
Ten Years of *Assessment Measures* 421

Acknowledgments

Many people helped to create this volume and each deserves special thanks. In the Preface I noted the significant contributions of my fellow columnists—Peter, Gary, Peter, and Jeff—Gale Erlandson at Jossey-Bass and Clara Roberts at NCHEMS. Without the work of these individuals, producing a retrospective volume drawn from *Assessment Update* would not have been possible.

Here I would like to express my deep appreciation to the others at IUPUI, Jossey-Bass, and NCHEMS whose cooperation, hard work, positive energy, and talents have brought the contents of the volume together.

Indiana University-Purdue University Indianapolis

- Karen Black
- Linda Durr
- Cheyenne Walker
- Emily Ward

Jossey-Bass Publishers

- Debra Hunter
- Vince Lewis
- Margie Macinenney
- Catherine Markham
- Sarah Miller

National Center for Higher Education Management Systems (NCHEMS)

- Paula Dressler
- Jay Reeves

Many thanks to all of you.

Trudy W. Banta

Editor's Notes

Trudy W. Banta

About the Author

Trudy W. Banta is vice chancellor for planning and institutional improvement and professor of higher education at Indiana University-Purdue University Indianapolis. Prior to assuming her current position in 1992, Banta was the director of the Center for Assessment Research and Development and professor of education at the University of Tennessee, Knoxville. She has developed and coordinated thirteen national conferences and ten international conferences on the topic of assessing quality in higher education. She has consulted with faculty and administrators in 36 states and Puerto Rico and has by invitation addressed national conferences in China, France, Germany, Canada, and Spain on outcomes assessment. Since 1983, Dr. Banta has edited six published volumes on assessment, contributed 19 chapters to published works, and written more than 80 articles and reports. She is the founding editor of *Assessment Update*, a bimonthly periodical published since 1989. Dr. Banta holds an Ed.D. in educational psychology from the University of Tennessee, Knoxville. Among her honors are recognition in 1988 for contributions to the field of assessment by the American Association for Higher Education, and recognition in 1997 by the American Productivity and Quality Center for leadership of one of six programs exemplifying best practice in using management information in decision-making.

Contents

Page

Reflections on Ten Years as Editor of *Assessment Update* 5

Section I
Encouraging Faculty Involvement in Assessment 7

On the Crest of the Wave 9
Can Assessment Help Research Universities? 11
Involving Faculty in Assessment 14
Faculty Involvement in Assessment: Why Is It So Hard to Achieve? 17
Teamwork Does Not Abridge Academic Freedom 22
Making a Silk Purse 25

Section II
The Growing Variety of Assessment Methods 29

Weaving Assessment into the Fabric of Higher Education 31
With FIPSE Support 33
TQM Without Assessment? How? 36
A Final Report to FIPSE on Assessment and Total Quality Management 39
An Audience of Beginners Needs Basics 42
Revealing the Results of Assessment 45
The Power of a Matrix 47
Where Are We in Assessment? This Issue Says It All 51
Terenzini's Purposes Frame This Issue 54
This One's for Students 57
The Responsive Community College 60
The Rising Star of Assessment Scholarship 63
Engineering Criteria 2000: Focus on Outcomes 66
Of Websites and Portfolios 69
Bob Pace Tells Us What Students *Do* While in College 72

Page

Are We Making a Difference? 76
Does Assessment Help Faculty Teach More Effectively? 79
Are We Making a Difference in Student Learning? 82
Now We Are Ten 85
Benchmarking in Assessment 87
That Summer Reading List 90

Section III

Assessment in Response to External Pressures:
Performance Funding in Tennessee 93

A Small Controversy 95
Some Unvarnished Truth 98
Performance Funding in Tennessee Comes of Age 100

Section IV

Some National Assessment Issues 103

A Community of Learners 105
Collaborating in Setting Directions for Assessment Research 107
Our Readers Speak (We Wish You'd Write!) 109
Education Indicators 111
Education Counts 113
The National Center on Postsecondary Teaching, Learning, and Assessment .. 116
Take Part in the National Goals Debate 118
Do Faculty Sense the Tightening of the Accountability Noose? 121
Soil to Support 1,000 Flowers 123
Can We Combine Peer Review and National Assessment? 126

Section V

International Perspectives on Assessing Quality in Higher Education 129

Assessment: A Global Phenomenon 131
Evaluation of Teaching and Learning in Germany 133
Assessing Quality—Toward Broader Perspectives 136
A Global Perspective—at Last 138
The Many Faces of Assessment Methodology 141
Further Steps Toward Globalizing Assessment 144

TRUDY W. BANTA

Reflections on Ten Years as Editor of *Assessment Update*

The first issue of *Assessment Update* was published in January 1989. For two years *Assessment Update* was a quarterly, but by mid-1990 the number of manuscripts had increased sufficiently to warrant more frequent publication and readers let us know that they would like to see more issues. In 1991 *Assessment Update* began to appear six times a year.

What I have learned in the course of preparing this retrospective is both gratifying and sobering.

To the 56 issues published thus far I have contributed 45 columns entitled *Editor's Notes*. While I have often looked back at individual columns to see what I had written previously about a topic I wanted to consider again, not until now have I re-read all the columns and contemplated them as a body of work. What I have learned in the course of preparing this retrospective is both gratifying and sobering.

In most of my columns I have referred in some way to the content of the articles and columns contributed to the issue by others. It is gratifying to see that the articles written by assessment coordinators and other faculty have become more interesting and sophisticated over time as more faculty have come to view writing about their experiences in assessment as legitimate scholarship in their own disciplines. What is sobering is how little my own range of topics and views of how assessment should be implemented have changed over the last decade.

As the individual with the responsibility for coordinating campus-wide assessment initiatives at first one, then a second, research-oriented university, most of my career has been based on finding ways to encourage faculty to become engaged in outcomes assessment and to improve their assessment methods continuously. Not surprisingly, then, it turns out that 30 percent of the columns I have written have addressed how difficult it is to get faculty to take on outcomes assessment and how to engage them, and 50 percent of my columns have focused on details surrounding one or more assessment methods. While I can see some progression in the development of assessment methods over the decade, the discouragement sets in when I see how little

BEST COPY AVAILABLE

difference ten years has made in terms of convincing faculty that outcomes assessment holds promise for making them more effective in everything they do.

In my first column, which appeared in January 1989, I wrote, "Most faculty, especially those in the nation's largest institutions, are just starting to hear the roar of the first waves of assessment hitting the shore." I used the sea metaphor because by 1989 I had already been leading a campus-wide assessment program at the University of Tennessee for a decade and thus had posed the question in my editorial, "Has the assessment wave crested?" While a lot of water has hit the shore in the decade since I wrote that first column, my observations about the state of affairs are the same today as they were then: most faculty may have heard the roar by now, but they are NOT ENGAGED in assessment, and thus the assessment wave cannot have crested-at least not in the way those of us associated with the production of *Assessment Update* might have hoped. Assessment cannot have reached its full potential as an accepted avenue toward improvement for all faculty when in 1998 assessment committee members at the largest research universities are assuring each other that outcomes assessment is "new" (most likely because THEY just discovered it!) and new presidents are "launching an assessment initiative" on campuses where standard-setting assessment programs flourished a decade ago but subsequently withered and faded away. Outcomes assessment apparently goes against the grain of those in the academy so profoundly that only extraordinary leadership can sustain it.

Most faculty, especially those in the nation's largest institutions, are just starting to hear the roar of the first waves of assessment hitting the shore.

Having vented my frustration at our slow progress in convincing faculty of the value of assessment, I should complete this introduction by saying that the 20 percent of my columns not focused directly on faculty involvement or descriptions of assessment methods have dealt with state-level, national, and international issues in outcomes assessment. In four columns I have described aspects of Tennessee's performance funding program, which first brought me into outcomes assessment in 1979. Five columns contain observations about National goal 6.5, which states that by the year 2000 "the proportion of college graduates who demonstrate an advanced ability to think critically, communicate effectively, and solve problems" will increase substantially. And since I began to co-sponsor an international conference on assessment the same year we inaugurated *Assessment Update*, eight columns have addressed international developments in the arena of assessing quality in higher education.

I have just noted the presence of five major themes in the series of *Editor's Notes* columns published in the first ten volumes of *Assessment Update*. These themes are used as section headings for the collection of columns that follows. Since many columns treat multiple topics, the numbers of selections in each section may not match the numbers cited above. Each section begins with a brief overview that attempts to place the contents in some perspective.

Section I.
Encouraging Faculty Involvement in Assessment

In the first selection (1:1) the growing pressures—regional, national, and international—to engage in outcomes assessment are cited as evidence that a national publication like *Assessment Update* is needed. The point is made, nevertheless, that most faculty are just beginning to become aware of the crescendo of calls for higher education to demonstrate its accountability through assessment of student learning and other outcomes society values.

Most faculty are just beginning to become aware of the crescendo of calls for higher education to demonstrate its accountability through assessment of student learning and other outcomes society values.

By 1992 (4:4) I had observed that faculty at research universities were among the last to take assessment seriously because their colleagues value research so highly that few feel they can spend precious time on improving their teaching through assessment. In 1993 (5:1) I argued that, "At its best, assessment can even help faculty do what they value in better ways and thus maximize their accomplishments." I noted that "the single most important outcome of faculty involvement in assessment may be the extent to which the goals and objectives for curricula and courses have been reconceptualized" as a result of collective engagement in the assessment process.

But by 1996 (8:5) I was back to square one explaining again why faculty involvement in assessment is so hard to achieve. And in 1997 (9:5) I felt compelled to address in detail one of the most troublesome assertions by reluctant academics; i.e., that agreeing with colleagues on a set of learning outcomes that should be taught in one or more courses may limit an individual teacher's academic freedom. I concede that this conclusion can emerge from a strict construction of the AAUP "Statement on Academic Freedom and Tenure." Nevertheless, I point out the advantages of teamwork and agreement upon some definitions of the common good for a group of colleagues.

For years those of us who care about assessment have been seeking ways to link it with one or more bodies of ongoing research in order to enhance its viability and credibility and to sustain its development. Thanks to a wonderful article by Peter Ewell that appeared in the December 1997 *AAHE Bulletin*, I was able to describe in 1998 (10:2) some of the areas of inquiry in which this linkage is beginning to occur.

BEST COPY AVAILABLE

TRUDY W. BANTA

On the Crest of the Wave

Response to the idea of publishing a newsletter on assessment in higher education has been immediate, strong, and positive. Jossey-Bass sent out a questionnaire designed to test the waters for a newsletter. Within a few weeks, over a third of the questionnaires came back, all supporting the concept, with most containing content suggestions and offers to contribute information.

We hope you will follow through with those offers. Since we don't have reporters strategically stationed around the globe, we will depend on our readers to send us the news. Articles of any length (up to 3,000 words) will be welcome. Send us announcements of meetings to come, reviews of meetings just held, word of collaborative efforts, insights about how assessment works (or doesn't work), stories about innovative assessment activities, and any other relevant information.

Gale Erlandson, higher education editor at Jossey-Bass Publishers, asked if I would serve as editor for a newsletter on assessment. It would be an entirely new venture for Jossey-Bass—its first newsletter—as well as for me, but both Gale and I felt this type of publication was exactly right for keeping up with a burgeoning and rapidly changing field. Always supportive administrators at UTK promised assistance. Longtime friends Peter Ewell, Pat Hutchings, Gary Pike, Marcia Mentkowski, and John Harris pledged their support. And more recent acquaintances Peter Gray and Larry Braskamp offered specific, immediate help. How could anyone say no in light of such enthusiasm? I promise that with your help and that of our consulting editors, we will try to report on the most promising practices and the most timely issues in assessment.

Having Peter Ewell as author of the lead article is a special treat. Since we met in 1982, Peter and I have been sharing our perceptions about assessment wherever we meet—in airports, at receptions, in deans' offices, speeding to the next appointment in a rented car. In 1986 we began to ask each other, "Has the assessment wave crested?" Each succeeding year we have admitted,

In 1986 we began to ask each other, "Has the assessment wave crested?" Each succeeding year we have admitted, with some incredulity, that interest has continued to grow.

Assessment Update
Spring 1989
Volume 1, Number 1

BEST COPY AVAILABLE

with some incredulity, that interest has continued to grow. Peter's theme "about halfway" expresses well the feeling many of us have about our position in the history of assessment in higher education.

Most faculty, especially those in the nation's largest institutions, are just starting to hear the roar of the first waves of assessment hitting the shore. Two-thirds of the states and all six regional accrediting agencies have taken some action to stimulate institutional interest in the assessment of outcomes. On July 1, 1988, the *Federal Register* published the final regulations for the Secretary of Education's "Procedures and Criteria for Recognition of Accrediting Agencies." The regulations include a section called "Focus on Educational Effectiveness," which specifies that each accrediting agency must determine whether or not an institution or program (1) maintains clearly specified educational objectives consistent with its mission; (2) documents the educational achievements of its students "in verifiable and consistent ways," such as evaluation of senior theses, standardized test results, and employer evaluations; (3) publicizes for the benefit of prospective students its educational objectives and the results of its assessment procedures; and (4) systematically applies the information obtained through assessment "to foster enhanced student achievement."

Interest in assessment is growing internationally.

The pervasive nature of the state mandates and accreditation criteria, as well as the genuine desire to promote student development, brought faculty all over the country together for state and regional meetings during the fall of 1988. At the annual meetings of the American Evaluation Association and the American Educational Research Association, sessions on assessment in higher education were added to the agendas; this indicates that assessment practitioners are building bridges to the research communities with which this fledgling field must be linked.

Interest in assessment is growing internationally. In 1986 UTK and Northeast Missouri State University were invited to make presentations in Paris at a meeting of the Organization for Economic Cooperation and Development. Faculty at Alverno College have made several presentations in England and Europe. In 1987 the first Sino-American Symposium on Program Evaluation in Higher Education was held at Peking University. In June, 1989, a second team of American assessment specialists traveled to Beijing for five days of presentations and dialogue with faculty from thirty or more Chinese universities. On July 24-27, 1989, UTK cosponsored an international conference on assessment held at Cambridge University. Conferees considered the feasibility of international collaboration on matters related to assessment.

We hope to keep you apprised of future developments in assessment through feature articles, short perspectives pieces, news items, a resource column, a calendar of coming events, and continuing features on state initiatives, campus programs, and assessment instruments. This launch issue was assembled several months prior to publication; however, more short pieces and more news are planned for future issues. If you have additional suggestions for the presentation of content, please send them to us.

Assessment Update
Spring 1989
Volume 1, Number 1

TRUDY W. BANTA

Can Assessment Help Research Universities?

A colleague in assessment in another state maintains that it was dissatisfaction with the undergraduate experience at research universities that originally provoked legislators in some states to require all public colleges and universities to gather evidence of student outcomes. He says, "Lawmakers are tired of hearing complaints from constituents about huge classes, incomprehensible international teaching assistants, and lack of attention from faculty that their children encounter on the state's flagship campus."

If my colleague is correct in saying that research universities were the first target of the state assessment mandates, it is ironic that, collectively, they have become the last institutions to provide meaningful responses. Usually in the forefront of scholarship and research activities, few Carnegie I universities have become noted for their faculty contributions to the literature on outcomes assessment.

Jeff Seybert, whose inaugural column on community college assessment is timely, welcome, and much appreciated, argues that two-year institutions have diverse missions; yet, it is hard to imagine more diversity than exists on a land-grant campus of 40,000 students that includes a medical school, a dairying program famous for its ice cream production, a business school that trains corporate managers, and a speech and hearing center that serves hearing-impaired clients throughout the region. Corralling the faculty in such a place to focus on any single issue is nearly impossible. The faculty tend to view state mandates as something administrators are paid to handle, and no one is seriously concerned that State U might lose its institutional accreditation if it fails to address the regional association's assessment criteria. In addition research-oriented faculty are the least likely to be rewarded for conscientiously evaluating and improving their teaching and the most skeptical about the technical adequacies of the measurement instruments currently used in assessment. Considering these realities, you begin to have some sense of the difficulty of mounting a campus-wide, comprehensive outcomes assessment program on a research-dominated campus.

Research-oriented faculty are the least likely to be rewarded for conscientiously evaluating and improving their teaching and the most skeptical about the technical adequacies of the measurement instruments currently used in assessment.

Assessment Update
July-August 1992
Volume 4, Number 4

BEST COPY AVAILABLE

But there are some important signals of a change in this state of affairs. In the last two years, faculty at institutions like Pennsylvania State, Michigan State, and the University of California, Berkeley, have completed studies indicating a need for major reforms of undergraduate education, and some see assessment as contributing to that effort. Teams of faculty from these and other research-oriented campuses are showing up at regional and national assessment conferences. Some have written grants with an assessment focus for funding from the Fund for the Improvement of Postsecondary Education and other granting agencies.

How can assessment help the faculty at research universities arrest the swing of the pendulum toward research and bring it back toward teaching and scholarship related to pedagogy?

Pat Hutchings, in her article in this issue, suggests that outcomes assessment activities can assist faculty in setting goals and creating curricular coherence. It is difficult to imagine how one might go about improving something without being able to state in writing what that something is. Yet most undergraduate programs of general education and the majors at research universities take place in the absence of faculty-developed statements of what students should know and be able to do when they complete a course of study. Many research faculty say there is no need to write anything down because "everyone knows" what should be taught. Supposedly, students will be properly prepared if they simply take a sampling of the courses offered by the strong faculty in the program (that is, the productive scholars). As faculty at the University of Connecticut and at Colorado State University have found, however, involvement in the hard work of developing a comprehensive outcomes assessment instrument often sensitizes academics to the need for structured goals for student learning.

How can assessment help the faculty at research universities arrest the swing of the pendulum toward research and bring it back toward teaching and scholarship related to pedagogy?

Similarly, faculty who develop assessment measures and use them to collect data soon begin to ask questions about curricular coherence. Several departments at the University of Tennessee, Knoxville, have been motivated by assessment findings to increase both the breadth and depth of the required curriculum in the major. At Western Michigan University, the assessment findings that too many students were failing courses for which they were not adequately prepared led to the decision to scrutinize students' prerequisite experiences much more carefully.

Student-faculty interaction is known to promote student persistence and growth during college. At a research university, student-faculty collaboration on research projects would seem to be a particularly appropriate mode of interaction. One of the most salient features of the Harvard Assessment Seminars has been the involvement of students in the design and conduct of data-gathering activities that have grown out of seminar discussions. At Indiana University, Bloomington, students have taken part in an assessment of how undergraduates study. At the University of Virginia, a longitudinal study designed to assess the effectiveness of general education has brought a sample of undergraduates into contact with a group of faculty for a series of well structured interviews.

Assessment Update
July-August 1992
Volume 4, Number 4

The Cross-Angelo classroom research techniques have been embraced by some research university faculty who are interested in formative evaluation of their own teaching. In large sections where few can ask questions of the lecturer, the one-minute paper can be of special value in identifying points that students need to have clarified.

Assessment can give faculty at research universities the data they need to direct faculty development resources, as well as institutional recognition and reward structures, toward the improvement of teaching and learning. Although these "new students" of assessment are late to class, the intellectual and creative contributions they bring to the discussion will be most welcome.

Assessment can give faculty at research universities the data they need to direct faculty development resources, as well as institutional recognition and reward structures, toward the improvement of teaching and learning.

Assessment Update
July-August 1992
Volume 4, Number 4

Involving Faculty in Assessment

In motivating faculty to take part in assessment, we must not present assessment as a new, independent activity but rather as one that is linked with work in which they are already engaged.

Assessment Update
January-February 1993
Volume 5, Number 1

In 1992 I was privileged to take part in two unique assessment conferences. The first, held in Seattle in early October, was sponsored by the National Center on Postsecondary Teaching, Learning, and Assessment (NCTLA) at the Pennsylvania State University. The second, the Assessment Workshops at Indianapolis, took place in early November on the Indiana University-Purdue University campus. The NCTLA Assessment Institute emphasized attendance by institutional teams and offered an intensive three-day program with several two-hour blocks for teams to engage in their own assessment planning, with assistance, if needed, from an experienced consultant. In Indianapolis, the emphasis was on interaction among participants in a series of workshops conducted by faculty from a variety of institutions.

During both conferences, I could not help thinking how much the audiences for assessment meetings have changed over the half-dozen years since the first one was held. Early participants were academic vice presidents and deans who were interested in learning how they could provide leadership for campus assessment programs. Now the meetings are dominated by faculty who are carrying out the campus assessment initiatives.

Although the nature of the audiences has changed, the fundamental questions raised about assessment have not. They are simply being asked with more intensity by the people who are most directly involved in implementing assessment.

One of the most critical questions has always been, "How do you motivate faculty to become involved in assessment?" In responding to this query in Seattle, I suggested that we look first at what faculty value. While the list of intrinsic and extrinsic faculty motivators is relatively short, the priorities among these factors vary considerably from campus to campus. I believe, however, that a common concern underlying all the factors is this: faculty value their time and want to use it in ways that maximize their accomplishments. Thus, in motivating faculty to take part in assessment, we must not present assessment as a new, independent activity but rather as one that is

BEST COPY AVAILABLE

linked with work in which they are already engaged. At its best, assessment can even help faculty do what they value in better ways and thus maximize their accomplishments.

At an early stage of their involvement, faculty can be encouraged to spend on assessment the time they might ordinarily spend on service to the institution, that is, in working on departmental or institutional committees. If an unusual amount of time is required for an individual to provide leadership for a program or for the design of a data-gathering activity, release time or extra-service pay might be offered; at a research university, a graduate student might be assigned to assist the faculty member who is assuming this level of responsibility. For extraordinary leadership, every effort should be made to provide special recognition, perhaps in the form of an award, perhaps through additional compensation.

In Seattle, I worked with a team from a technical institute where participation in professional development experiences is the only avenue to differential increases in compensation. We talked about treating assessment, with its emphasis on new ways of collaborating to improve curriculum and instruction and its workshops on objective writing and test construction, as faculty development. As we were having that conversation in Seattle, a campus-wide faculty meeting on assessment was underway at the technical institute. Warming to the idea of assessment as professional development, the team members decided to telephone their dean for instructions and to ask him to announce that participation in assessment work groups scheduled for the afternoon at the institute would carry with it professional development credit. The dean liked the idea and agreed to make the announcement. Seldom have I seen a new idea implemented so quickly!

As more and more faculty become engaged in systematic efforts to improve processes in higher education, I believe that they will come to regard assessment as an indispensable resource for those activities.

On research-oriented campuses, involvement in assessment can be construed as time spent improving one's teaching and can even become the subject of original scholarship and research. Participation in the Harvard Assessment Seminars has led a number of faculty, working with undergraduate and graduate students, to design a series of studies to provide answers to questions about teaching and learning. Across the country, faculty in a broad array of fields have begun to publish articles about assessment activities in journals in their own disciplines.

As more and more faculty become engaged in systematic efforts to improve processes in higher education, I believe that they will come to regard assessment as an indispensable resource for those activities. At the Indianapolis conference, we speculated that the single most important outcome of faculty involvement in assessment may be the extent to which the goals and objectives for curricula and courses have been reconceptualized. Instead of the content that faculty will put before students, objectives increasingly emphasize what students will be able to do with that content. Greater clarity about what we want to accomplish is a first step in process improvement. Equally important is collaboration with colleagues in studying the details of the process, another *sine qua non* of assessment. Finally, the assessment

Assessment Update
January-February 1995
Volume 5, Number 1

measures that we develop are essential tools in the monitoring phase of a continuous improvement initiative.

Although there is no single or simple answer to the question of how to get faculty involved in assessment, since the response varies from campus to campus depending on the basis for faculty recognition and rewards, we still can and must make assessment an integral part of the activities that faculty value. Then they can make their time go farther by doing two (or more) things at once: assessment and committee service, professional development, disciplinary scholarship and research, and continuous improvement of teaching and learning.

Although there is no single or simple answer to the question of how to get faculty involved in assessment, since the response varies from campus to campus depending on the basis for faculty recognition and rewards, we still can and must make assessment an integral part of the activities that faculty value.

Assessment Update
January-February 1993
Volume 5, Number 1

TRUDY W. BANTA

Faculty Involvement in Assessment: Why Is It So Hard to Achieve?

Faculty involvement in assessment is the theme of this issue of *Assessment Update*. Virtually every article touches on one or more aspects of the questions I raise in this column.

John A. Muffo devotes a whole section of his "Lessons Learned from a Decade of Assessment" to faculty concerns. Charles F. Harrington and Robert L. Reid begin their article with the assertion that "one of the more pressing issues facing the assessment movement in higher education is that of engaging the faculty." And in one of the Memos, Sheila Bassoppo-Moyo observes that faculty at her institution are not likely to get involved in assessment without some external impetus.

We can all identify reasons for faculty resistance to assessment.

In a previous issue of *Assessment Update* (Vol. 8, No. 3), Peter Ewell reported on a recent survey of state higher education executive officers in the 50 states. His respondents identified resistance to assessment among faculty and administrators as the most important obstacle to its progress.

In the March-April 1996 issue of *Assessment Update* (Vol. 8, No. 2), Joe Steele summarized the findings of research conducted with a national sample of college and university representatives. According to Steele, "Results for the majority of assessment programs have been disappointing" (p. 2) and "the lack of faculty involvement in assessment activities" is one of the "two most critical outcomes issues specifically mentioned by more than half of the responding colleges" (p. 12).

Why Do Faculty Resist Assessment?

We can all identify reasons for faculty resistance to assessment. When first presented with the possibility of developing an approach to the assessment of student outcomes, most faculty respond, "We assess all the time. We use course and cumulative grade point averages to tell us how students are doing. What's wrong with that?" We might respond, "Yes, you do a

Assessment Update
September-October 1996
Volume 8, Number 5

BEST COPY AVAILABLE

lot of assessment of individual student achievement, and that is the most immediate way to help students. But with our new emphasis on *outcomes* assessment, we need to look carefully at the achievement of *groups* of students to see what that may tell us about how to enrich our teaching, enable students to learn more, and improve the curriculum." The faculty response is often, "Why would we want to change the curriculum? Things are working well now."

Most faculty eventually admit, even though it may be grudgingly, that nothing is so perfect that it cannot be improved at least a little. So they ask the next question, "Where will we find the time to take on this new approach to assessment? And where will our departments get the money it will take to do this new work?" Faculty also express concern that collective agreement with colleagues on objectives for a curriculum will have at least two negative consequences: (1) students then will be led to certain conclusions rather than forced to put their coursework together mentally and make sense of it on their own, and (2) each individual faculty member's academic freedom will be abridged by having to devote some of his or her teaching efforts toward promoting students' learning of the common curriculum objectives.

The request to take time to learn new methods designed to improve pedagogy is often resented. Most faculty are not trained as teachers and many believe that good teachers are born not made. Thus teaching is considered a private matter that is not often discussed extensively with colleagues. Faculty tend to shun teacher training and feel there is no need for them to engage in it.

Finally, many faculty, especially those at research universities, consider their scholarship—the creation of new knowledge—to be more important and more interesting than classroom teaching. Thus time spent on improving teaching is perceived as detrimental to the pursuit of the more important disciplinary scholarship.

But with our new emphasis on outcomes assessment, we need to look carefully at the achievement of groups of students to see what that may tell us about how to enrich our teaching, enable students to learn more, and improve the curriculum.

Can Faculty Continue to Ignore Assessment?

Some of our colleagues are saying, "The older generation of faculty is simply going to have to leave in order for us to make progress in turning attention toward student learning and the improvement of teaching." I don't agree with this because I know too many 50- and 60-year-olds who are enthusiastic proponents of assessment. And yet, even if those who make this prediction are right, in some places the older generation is leaving faster than any of us would have thought possible five years ago. In Canada and in Australia, just to name two countries with which I have had very recent experience, government cuts in funds to support higher education are so substantial that many faculty aged 55 and older are being offered attractive severance packages, which most feel compelled to accept. These actions, based on economic factors, are wiping out a whole generation of scholars in one fell swoop.

Assessment Update
September-October 1996
Volume 8, Number 5

TRUDY W. BANTA

The times are changing—more rapidly than ever before. If there ever was a time when faculty could legitimately say, "We can ignore assessment; it will go away," that time has certainly passed by now. If one has any doubt that entities such as higher education institutions, which exist to provide services to others, can escape the obligation to demonstrate their accountability, one has only to look at the enormous changes in health care that are occurring at lightning speed and at the outcomes-based agenda of every publicly supported social services agency.

There are some who say that competition will shortly be responsible for the most dramatic change in higher education. Certainly, if student numbers fall, so eventually must the number of faculty. And the number of sources from which potential students may choose their method of obtaining postsecondary education is increasing rapidly.

The program of the Eighteenth Annual Forum of the European Association for Institutional Research (1996), held in Budapest in August, included a session on "Alternatives to Universities." The developments considered in the session were "related to the growing market orientation of higher education systems in many parts of the world."

There must be assessment of student competence—evidence that students are mastering the course and curriculum objectives.

In business and industry employers want their employees to develop particular skills and understandings. Rather than deal with colleges and universities whose broader-based courses may or may not deliver the level of specific skills desired, employers are establishing their own courses or contracting with private firms to develop the skills in question. Even public schools that used to look to colleges and universities for in-service training for teachers and supervisors are now likely to ask their own expert practitioners or to contract with a training firm to develop the needed skills.

Taxpayers and their representatives are looking for ways to educate the populace at the lowest possible cost. The governors of eleven Western states are discussing the establishment of a virtual university that would award credit by assessing students' abilities to demonstrate competence in college subjects. Coupled with interactive tele-learning that can make one well constructed course available to students anywhere in the world, the virtual university concept has the potential to change fundamentally the way students experience postsecondary education.

All of these developments taken together suggest that faculty must become more intentional about what they will teach and what students will learn. There must be explicit, agreed-on objectives for curricula and for the courses that comprise them. Above all, there must be assessment of student competence—evidence that students are mastering the course and curriculum objectives. Given the ferocious competition developing among learning organizations worldwide, these are simply necessary, but not sufficient, steps that faculty must take to preserve their own jobs and perhaps even the very existence of their institutions.

◆

Assessment Update
September-October 1996
Volume 8, Number 5

What Can We Do?

Do those of us who are concerned about assessment simply wait until these outside pressures force faculty to become involved? That doesn't seem prudent. We must take advantage of opportunities to make assessment a part of what faculty value: disciplinary accreditation, evaluation of pedagogical and curriculum innovations, improvement of teaching, faculty development, program review, continuous quality improvement, and evaluation of technology in instruction. Not every one of these connections will have importance for every faculty member, but perhaps one will capture the attention of each.

We should encourage participation in faculty development such as trips to conferences, on-campus visits by experts, and homegrown seminars. We should offer small grants for innovative work and give faculty abundant opportunities to learn from one another. As Ann Ferren (1993) observed about faculty development, "To make assessment real for faculty, they have to experience it, not just read about it or hear about it. Campus-based training for assessment can extend faculty understanding and commitment and eliminate fears about what will happen to assessment data" (p. 7).

We must take advantage of opportunities to make assessment a part of what faculty value.

We can involve students in assessment. They will ask the hard questions that show faculty why they must undertake assessment, especially at an urban campus or a community college. We can encourage faculty to work in the community, where they will hear from employers, public school people, parents, taxpayers, and social service agencies that we must provide credible evidence of what our students are learning.

Many of the suggestions just made about involving faculty are richly illustrated in the contributions to this issue. Some offer additional ideas. Muffo offers specific strategies for overcoming faculty resistance to assessment based on his decade of work in this field. Harrington and Reid as well as Donald H. Bennion and Stewart D. Work provide details of the process of establishing faculty committees and involving faculty in the selection and design of assessment instruments. Elizabeth Domholdt describes an innovative approach to faculty development: inviting individual faculty to develop posters about their work in assessment for presentation at the annual Faculty/Staff Institute at the University of Indianapolis. In Barbara S. Fuhrmann's Campus Strategies, she lists among the success factors strong administrative support in the form of release time for assessment activities. And, finally, Daniel J. Bernstein's FIPSE project holds promise for drawing faculty attention to assessment by incorporating in peer review of teaching a focus on student learning and faculty feedback to students about their performance.

Assessment Update
September-October 1996
Volume 8, Number 5

Faculty response to the calls for assessment is just one of many ways in which the mismatch between what society is beginning to expect and what higher education is actually delivering is becoming manifest. It is simply one illustration of how faculty attitudes must change in order to ensure the survival of higher education as the enterprise we know and value. For those interested in a broader look at these issues, I strongly recommend *Making* *a Place for the New American Scholar* by Eugene Rice (1996), director of the AAHE Forum on Faculty Roles and Rewards. In addition to calling for important changes in the graduate training we provide for prospective faculty, Rice offers a series of suggestions for "rethinking faculty careers." These include creating a departmental culture in which (1) collaboration is valued, assessed, and rewarded, and where faculty can easily cross disciplinary lines to collaborate with colleagues in other fields; (2) excellence in teaching, instructional scholarship, and public service are as valued and rewarded as research is today and faculty have the opportunity to concentrate on each of these areas at different points in their careers; and (3) faculty have opportunities to spend extended periods of time learning from practice in settings outside the academy—making professional contributions of a different kind—then moving back their academic appointments. Rice also suggests that tenure be considered local, tied to institutional mission. Thus, if assessment of student outcomes and institutional effectiveness are valued sufficiently to be incorporated in an institution's mission, its tenure-track faculty will have a powerful incentive indeed for becoming involved in assessment!

If assessment of student outcomes and institutional effectiveness are valued sufficiently to be incorporated in an institution's mission, its tenure-track faculty will have a powerful incentive indeed for becoming involved in assessment!

References

Eighteenth Annual Forum of the European Association for Institutional Research. "Higher Education in the Market Place: Strategies of Survival and Success." Preliminary program, August 25-28, 1996, Budapest, Hungary.

Ferren, A. "Faculty Resistance to Assessment: A Matter of Priorities and Perceptions." In *Conference Discussion Papers of the Eighth Annual Assessment Conference and the First Continuous Quality Improvement Conference*. Washington, D.C.: American Association for Higher Education, 1993.

Rice, R. E. *Making a Place for the New American Scholar*. Washington, D.C.: American Association for Higher Education, 1996.

Assessment Update
September-October 1996
Volume 8, Number 5

Teamwork Does Not Abridge Academic Freedom

True teamwork in the academy is not well developed.

Teamwork may be the most profoundly influential word of the past two decades. Giving everyone who has responsibility for getting a job done the opportunity to say how that job can be accomplished most effectively has led to flatter organizations and more effective outcomes throughout the economy.

Teamwork in the fullest sense of the word has brought the American automotive industry back to a position of international leadership and enabled scientists and engineers to launch a spectacular geology expedition on Mars. Doctors, nurses, and other health care professionals approach surgery in well coordinated teams. And attorneys work in teams to gather evidence and plan arguments in complex cases.

All around us we see proof that even the brightest idea that springs from the fertile mind of a creative individual can be brought to fruition most effectively with the help of other good minds thinking through the how-to-get-it-done steps. But despite this compelling evidence, true teamwork in the academy is not well developed.

We understand the reasons for this. Most of the senior faculty and administrative leaders have worked in isolation for most of their careers—spending many hours studying alone to make good grades as undergraduates, then working on individualized programs and research projects in graduate school, often in a highly competitive environment that discouraged collaboration. Now promotion and tenure committees cast bright smiles on pure disciplinary scholarship but discount multidisciplinary activities and, in some fields, co-authored publications. Even authentic teaming to teach within a given discipline is a rarity.

This tendency to function as individual, independent scholars works against the acceptance of outcomes assessment in higher education. The essential first step in planning assessment in the major or in general education is for groups of faculty (dare we say "teams?") to get together and agree

Assessment Update
September-October 1997
Volume 9, Number 5

BEST COPY AVAILABLE

on learning outcomes that students who complete a program should achieve. Many of our colleagues resist this first step. A common argument, often stated passionately, is that having someone, even one's own colleagues, tell an instructor what a student should learn in his or her classes is an abridgment of academic freedom.

Having confronted this argument in settings across the country, I decided to consult the American Association of University Professors (AAUP) "Statement on Academic Freedom and Tenure" (see <http://www.igc.apc.org/aaup/1940stat.htm>). Now that I have read this simple statement—"Teachers are entitled to freedom in the classroom in discussing their subjects"—I understand that one can interpret it as placing no restrictions on what is taught beyond "controversial matter which has no relation to [the] subject" (the only qualifying phrase in the AAUP statement). This means that we as faculty must discuss and decide if there is some common good that we wish to advance in terms of student learning. And if there is, can we agree that we can rely on one another to help students develop certain skills and knowledge in each of our classes? Only if we can reach agreement on these issues can we hope to establish the foundation for outcomes assessment.

As we attempt to persuade new colleagues in assessment to work in teams to develop agreement about what students need to know and be able to do when they complete a program of study, we are also fortunate to be able to point to the benefits of such teamwork at other institutions.

Fortunately, for those of us with a firm commitment to assessment, the AAUP statement contains other, helpful language, including "institutions of higher education are conducted for the common good and not to further the interest of either the individual teacher or the institution as a whole. The common good depends upon the free search for truth and its free exposition." Moreover, "Academic freedom . . . carries with it duties correlative with rights."

As we attempt to persuade new colleagues in assessment to work in teams to develop agreement about what students need to know and be able to do when they complete a program of study, we are also fortunate to be able to point to the benefits of such teamwork at other institutions. Articles in this issue of *Assessment Update* provide examples.

At Pepperdine University, faculty from the humanities, sciences, and social sciences are collaborating to assess student achievement of the Pepperdine mission component stating that "all graduates will have the ability to think clearly, logically, independently, and critically." As Don Thompson and Cynthia Cornell Novak report, the faculty ask students to submit portfolios containing course syllabi, tests, papers, projects, and videotaped speeches. Portfolio materials are digitized and recorded on CD-ROM disks that students can use in job searches. The portfolios are also available on Web pages that faculty can study in their own offices. Pepperdine faculty have agreed to have colleagues judge their own syllabi and grading practices in terms of the extent to which their course content and instructional methods promote student achievement of clear, logical, independent, critical thinking. For these faculty, collaboration in identifying common expectations for student learning and in assessing outcomes does not abridge academic freedom.

Assessment Update
September-October 1997
Volume 9, Number 5

At the University of Scranton, Thomas P. Hogan and Anne Marie Stamford report that department chairs are asked not only "What are you trying to accomplish with your students?" but also the assessment question "Is there evidence you're being successful?" Underlying these questions is the clear assumption that faculty in the disciplines have collectively identified common learning outcomes for their majors.

Lynne M. Woehrle reports that "team building through regular team meetings" of sociology faculty at Syracuse University has led to adoption of a core set of readings and many of the same assignments in sessions of introductory courses taught by different faculty. Team members even visit one another's classrooms. While Woehrle acknowledges that there are "tensions between the pure academic freedom model, wherein each instructor defines course content, and the team model, wherein goals are shared," collaboration in assessment has "proved crucial in maintaining the internal financial support of the administration" for the course renovation project in sociology.

Collaboration in assessment has "proved crucial in maintaining the internal financial support of the administration."

Finally, in his *Campus Profiles* column, Peter J. Gray tells the story of assessment's influence at Northeast Missouri State University—now Truman State University. Through three decades, assessment has been at the center of that institution's mission and campus culture. The president and other top administrators have communicated the importance of assessment in "personal, one-on-one conversations with faculty members about specific students and their success at the university." In turn, faculty understand and appreciate—and communicate to students—that assessment is an important element of the common good, the culture that is shared by faculty, staff, and students. At Truman State, the "free search for truth and its free exposition" are not hindered but in fact are strengthened by teamwork based on shared purpose and the free exchange of information about progress that is derived from a conscientiously implemented assessment process.

Assessment Update
September-October 1997
Volume 9, Number 5

TRUDY W. BANTA

Making a Silk Purse

When you read in this issue's lead article the responses to the survey that Andreea Serban and her colleagues at the Rockefeller Institute have conducted, you will find some familiar themes. Compared with state and university policymakers, "deans and chairs of faculty senates are the least familiar with performance funding practices both within their own state and in other states." "State policymakers tend to forecast a favorable future for performance funding, whereas campus representatives think its future is uncertain." All campus groups "consider external accountability to be currently the main purpose of performance funding . . . [but] wish to see institutional improvement become the primary purpose." And "state policymakers . . . are stronger believers in the potential of performance funding to achieve its goals than are campus representatives."

Most faculty don't take the time to become familiar with the details of externally imposed accountability measures like performance funding because they rather resent them and believe they will eventually go away. Whereas state officials generally hope and believe that performance funding will improve higher education in the long run, faculty and administrators see such measures as intrusive and prefer to rely on their own internal methods for assuring quality.

> *Most faculty don't take the time to become familiar with the details of externally imposed accountability measures like performance funding because they rather resent them and believe they will eventually go away.*

Serban's article brings back memories of University of Tennessee, Knoxville chancellor Jack Reese's discomfiture with performance funding as conceptualized by the Tennessee Higher Education Commission (THEC) in 1979. He said, "How can we convince faculty that all the tests and statewide surveys the THEC is proposing are worthwhile?" As Reese's administrative intern I was delighted to find in my mail a few months later an opportunity to apply for a grant "to increase the use of student outcome information" offered by the Kellogg Foundation through the National Center for Higher Education Management Systems (NCHEMS). We applied for and received a grant, and over the next couple of years we were able to involve faculty in pilot testing some of the state-specified exams and surveys. Thus we gathered the student outcome information called for by the grant, but for our

Assessment Update
March-April 1998
Volume 10, Number 2

BEST COPY AVAILABLE

own reasons (improvement of programs and services) and under the banner of the "NCHEMS-Kellogg Project" rather than the "THEC's Performance Funding Initiative." From a faculty perspective, we tried, as my grandmother would say, to make a silk purse of a sow's ear.

Today, nearly two decades later, Butler University in Indianapolis is using a grant from the Lilly Endowment to create an opportunity for faculty to approach assessment as a means of achieving their own purposes. The Lilly-sponsored "Learning Initiative" at Butler is intended to "focus campus dialogue toward carefully defining Butler learning goals and objectives," then to "design effective learning activities with students using the best known educational practices" and "assist students in demonstrating their learning relative to Butler's learning goals" (quoted from *The Learning Initiative at Butler University* project description). As I sat at a dinner table at Butler recently between President Geoffrey Bannister and Dean of Academic Affairs Gwen Fountain and listened to them enthusiastically endorsing the goals of the Learning Initiative, I remembered the similarly high hopes with which Chancellor Reese and I had launched the NCHEMS-Kellogg Project. The next day as I met with small groups of Butler faculty to discuss their departmental assessment plans, I was pleased to see how well the project is succeeding in opening even some previously closed minds to the possibilities of using assessment to help strengthen a true learning environment at Butler.

Many of us have tried to find means of encouraging faculty to assume ownership of assessment, because our intuition told us it was the right thing to do.

Many of us have tried to find means of encouraging faculty to assume ownership of assessment, because our intuition told us it was the right thing to do. Now in a wonderful article in the December 1997 *AAHE Bulletin* entitled "Organizing for Learning—a New Imperative," Peter T. Ewell cites findings from the literature on organizational change that give us a foundation for our intuitive actions.

First, despite all the negative feelings that externally imposed measures like performance funding arouse in faculty, the fact is that these measures do stimulate institutional change that probably would not have occurred without them. Ewell says that "change requires a visible triggering opportunity," and "part of the art of transformational leadership is to recognize and capitalize on such opportunities when they arise" (p. 6).

Second, "change requires people to relearn their own roles." According to Ewell, faculty must be imbued "with a sense of collective accountability for learning of the same character and depth as is currently accorded scholarly research" (p. 6). Helping to create this sense of accountability was my most important responsibility as a faculty development consultant at Butler.

Finally, studies tell us that "change requires conscious and consistent leadership" (p. 6). Academic leaders must become "leading learners, . . . creating new lines of lateral communication and alternative reward structures," and paying attention to people's "feelings, perceptions, and symbols" (p. 6).

Assessment Update
March-April 1998
Volume 10, Number 2

The individuals who created the New England Educational Assessment Network, described in the article in *Assessment Update* (Vol. 10, No. 2) by Jean Woodbury, Bernard Gill, and Susanne Tracy, are applying the principles Ewell cites as they assess their colleagues' needs and provide responsive developmental experiences.

To my surprise, the lessons about stimulating faculty involvement in assessment were learned late at Ohio University. I first met Michael Williford (see Peter J. Gray's column in *Assessment Update* Vol. 10, No. 2) in the early 1980s when we both won NCHEMS-Kellogg grants. I admired very much his survey program and other methods for gathering student information. Therefore, after years of writing and speaking about Ohio University as an exemplar of good practice in assessment, I was somewhat shocked to learn that the North Central Association (NCA) visiting team expressed strong concerns about that institution's assessment efforts. The central administration had designed the plan, collected the data, and used the results, but the visitors did not find faculty sufficiently engaged in the process. The NCA directed Ohio University to refocus its assessment program on "department-based assessment to improve teaching, learning, and student services."

No one will be surprised to learn that Ohio University faculty expressed "some resentment about developing an assessment plan for the NCA!" But the process of providing faculty development and technical support that Williford has initiated should eventually produce the sense of collective accountability for student learning that will help to transform the proverbial sow's ear.

The process of providing faculty development and technical support . . . Williford has initiated should eventually produce the sense of collective accountability for student learning that will help to transform the proverbial sow's ear.

Reference

Ewell, P. T. "Organizing for Learning—a New Imperative." *AAHE Bulletin*, 1997, 58 (4), 3-6.

Section II. The Growing Variety of Assessment Methods

As I have traveled around the world I have noted that peer review is the most universally respected method of program assessment. When valued colleagues are brought in for a close inspection of an academic unit and asked for judgments about current program quality and future directions, it seems logical to present in a self-study for the visitors' review information about program resources, processes such as curricula and instructional methods, and outcomes. I have presented the case for incorporating assessment findings as a source of information in program review in several of my columns, including 1:2, 3:3, and 3:4 in this section.

As I have traveled around the world I have noted that peer review is the most universally respected method of program assessment.

Throughout its quarter-century history, the Fund for the Improvement of Postsecondary Education (FIPSE) has been an important and consistent source of funding for assessment projects. I paid tribute to FIPSE in 3:1 as we launched a continuing feature, *With FIPSE Support*, which summarized FIPSE-supported assessment projects until space pressures brought on by a new format forced us to discontinue the series in 1997.

In 3:1 I also spoke for the first time about the logical and mutually beneficial links that exist between outcomes assessment and total quality management (TQM). In 1992 I received FIPSE funding to translate the principles of TQM from the language of business and industry into that of higher education. That project is summarized in 4:3 and 5:2.

I have long been interested in finding out what works in assessment, and was pleased to report in 3:1, p. 7 on the development of a document entitled *Principles of Good Practice for Assessing Student Learning*. Setting objectives, selecting or designing instruments, collecting data, disseminating results, and following up to ensure that findings are used are some of the steps involved in good assessment practice, and are covered in 6:3, 3:3, and 8:4. Note particularly the stratagem of using a matrix to communicate, check, and track assessment processes (8:4).

BEST COPY AVAILABLE

By 1998 (10:4) I was ready to describe benchmarking as an assessment tool. Some first-hand experience with the methodologies employed by the American Productivity & Quality Center in Houston has shaped my concept of benchmarking as a process of standard-setting-looking for exemplary practice in a well defined area, investigating carefully the components of such practice, distilling some generalizations, and trying to apply that learning to improve one's own practice.

As early as 1991 (3:4) I included in a column a generalization that sounds as solid today as it did then: ". . . faculty at two- and four-year institutions are using a wide variety of assessment techniques, ranging from commercially available survey instruments to faculty-developed written assignments, portfolios, and interviews . . . the variety of faculty initiatives is growing faster than is the adoption of commercial methodologies." The rich variety of methods and resources described in my *Editor's Notes* emanates primarily from the work of contributors of feature articles and columns by others in *Assessment Update*. A partial listing of methods and the columns in which they may be found appears below:

Benchmarking 10:4
Classroom assessment 5:3
Community college strategies 7:5
Focus groups 3:4 and 7:5
Performance indicators 7:5
Portfolios 3:4, 8:6, 9:2, and 9:3
Primary Trait Analysis 9:3 and 10:5
Projects (senior design) 9:3
Self assessment 3:4
Surveys 3:4, 8:6, and 9:3
Syllabus analysis 8:6
TQM/CQI 5:3, 7:3, and 7:5
Tracking 7:5
Websites 9:2

By 1998 I was ready to describe benchmarking as an assessment tool.

Note particularly the column in issue 7:3. I am especially proud of the issue by and about students and the many helpful ways they can be involved in assessing outcomes.

I am also proud of my long association with C. Robert Pace, venerable developer of the College Student Experiences Questionnaire and the Community College Student Experiences Questionnaire. See 8:1 for illustrations of the use of the CSEQ and the CCSEQ in campus assessment programs.

Finally, my columns 5:5, 5:6, and 6:2 are based on my book *Making a Difference: Outcomes of a Decade of Assessment in Higher Education* (1993) San Francisco: Jossey-Bass. I drew upon experiences others had contributed to that book to provide answers to three questions about the impact of assessment over the decade of the 80s:

Are students learning more? (6:2)
Are faculty teaching more effectively? (5:6)
Are colleges and universities improved? (5:5)

TRUDY W. BANTA

Weaving Assessment into the Fabric of Higher Education

With most people in higher education now aware of assessment, the movement has entered an intensive implementation phase. Experimentation with a variety of methods is taking place. A few institutions have moved with their assessment efforts into the phase of the change process known as *institutionalization*. Assessment leaders with programs at this stage of development are arguing strongly that assessment should be incorporated into such established institutional practices as planning, student development programming, and comprehensive academic program review.

Experimentation with a variety of assessment methods is taking place.

Critics have expressed the hope that assessment will not command as much attention in the future as it has in the past four years, because it has "overshadowed many of the other critical issues facing higher education." I believe this view implies the failure to recognize the positive role that assessment can play in helping to achieve these and other important goals, if it is firmly woven into the institutional tapestry.

At pioneering institutions like Alverno College and Northeast Missouri State University, early assessment initiatives created unique campus identities that began to attract outstanding students and faculty. Since then, Alverno has based an entire undergraduate curriculum on closely monitored development of individual students, an approach that has brought opportunities for external funding and for sharing methods with other institutions that hope to duplicate Alverno's successes (see Stone and Meyer, 1989, in *Resources*). In its second decade of assessment, Northeast used a sophisticated knowledge base, derived from assessment, to plan its transformation from one of several regional universities into Missouri's only liberal arts institution.

In 1983, on my own campus, assessment was made a component of a strong nine-year-old program review process. Recently, one of the University of Tennessee, Knoxville, humanities departments produced an exemplary self-study, which made good use of data we provide from surveys of faculty, undergraduates, and graduate students, information that our assessment office provides. The self-study process fostered consensus among the faculty

Assessment Update
Summer 1989
Volume 1, Number 2

BEST COPY AVAILABLE

about program strengths and weaknesses and future directions for the department. External reviewers were able to reach conclusions early in their campus visit and then to concentrate on the details of making recommendations for departmental improvements that will be considered and acted on by the faculty and by the departmental, college, and university administrators.

Colleges that engage in rational planning—specifying objectives for student development; implementing curricula, methods of instruction, and campus services to accomplish these ends with students; and using assessment results to chart progress and make warranted improvements—can secure the future of assessment, because it will be such an integral part of our day-to-day activities. Moreover, assessment will not be viewed as a competing demand for attention if it is considered a *means to effect improvement* in every institutional process.

Colleges that engage in rational planning . . . can secure the future of assessment.

Assessment Update
Summer 1989
Volume 1, Number 2

TRUDY W. BANTA

With FIPSE Support . . .

The majestic Mayflower Hotel in Washington, D.C., was the site in October 1990 of the project directors' meeting convened annually by the Fund for the Improvement of Postsecondary Education (FIPSE). I became more aware than ever this year that this hardy band of faculty and administrators from institutions across the country truly represents "the best and brightest" among those who are passionately concerned about improving the process of teaching and learning in higher education. The usual quiet of the dark wood-paneled corridors in the hotel's meeting space was dispelled by the clusters of academics excitedly discussing with project creators a brave new approach to teaching introductory physics, a series of videotapes designed to serve as a faculty- and staff-development tool in combating racism on campus, and an interactive computerized test of reflective judgment.

Only the most savvy and determined of academic entrepreneurs can be successful in marshaling the campus and external support needed to win the award of FIPSE funds.

These projects began, almost a year before they were funded, as carefully honed five-page descriptions. They survived reviews, by multiple critics within and outside FIPSE, that cut nearly 2,000 entries to about 200 that were considered strong enough for fuller development in 20-page proposals. In preparing the detailed proposals, the project creators had to convince decisionmakers in their own settings to provide a large measure of the resources that would be needed for implementation. Next, these projects survived a second series of reviews that reduced their number to about 75. Finally, they weathered the intense scrutiny of budget negotiators. Only the most savvy and determined of academic entrepreneurs can be successful in marshaling the campus and external support needed to win the award of FIPSE funds.

As Connie Cook, former FIPSE program officer, pointed out in the Summer 1989 *Assessment Update*, the fund has been supporting projects on assessment of student learning outcomes since its inception in 1972. In the 1970s, the focus was on assessment of *individuals*, as exemplified in competency-based testing projects. By 1986, the emphasis had shifted to the assessment of *programs*, and FIPSE funded two centers of support for the development of this direction: the Assessment Forum, which sponsors the largest annual

Assessment Update
January-February 1991
Volume 3, Number 1

BEST COPY AVAILABLE

national conference, and the Assessment Resource Center (now the Center for Assessment Research and Development) at the University of Tennessee, Knoxville.

Awed by the array of talent assembled by FIPSE at the Mayflower, I was delighted that new program officer Sherrin Marshall convened an "assessment cluster" on the first afternoon of the conference. Most of the 29 assessment projects in the current FIPSE inventory were represented. What a wonderful opportunity to review the state of the art in our field!

I learned that the National Governor's Association and the Education Commission of the States received grants to provide technical assistance to states, with the aim of blending politicians' interest in accountability with higher education's interest in improvement. The Council on Postsecondary Accreditation is attempting to explicate the role of outcomes assessment in the accreditation process, while the Southern Association of Colleges and Schools is developing a training program for individuals serving on accreditation teams. The American Association of Community and Junior Colleges is working on student-tracking systems as a source of data for assessing institutional effectiveness. Three institutions—Benedict College, Rhode Island College, and Indiana University/Purdue University at Indianapolis—are attempting to improve measures of teaching effectiveness in order to strengthen the emphasis placed on teaching in faculty-reward systems. Winthrop College is modeling a statewide consortium approach to meeting a legislative mandate for assessment.

Several institutions are experimenting with assessment in at least one discipline.

Several institutions are experimenting with assessment in at least one discipline. Brown University and the University of Maryland, Baltimore County, are concentrating on foreign languages. The National Women's Studies Association has its own project. The University of Missouri School of Law is evaluating the use of videotapes in law instruction. At Austin Peay State University, a national cooperative test-development project is underway. Faculty at several institutions completed surveys identifying content area emphases in psychology, social work, political science, and English literature and then contributed multiple-choice items for item pools in each content area. Items are drawn from these pools to make up comprehensive disciplinary exams that can be administered at each of the participating institutions.

At the Texas College of Osteopathic Medicine, expert systems are being employed to assess and improve the problem-solving abilities of medical students. FIPSE support has assisted the University of Tennessee, Knoxville, to develop an employer survey and experiment with the strategic quality-management concepts described in Daniel Seymour's article on page one of this issue.

FIPSE has funded no fewer than a dozen projects that focus on improving assessment of general education (or liberal studies) outcomes. Portfolios, freshman and senior essays, classroom research methods, a "general education profile" for each student that can be used in advising, course-embedded

Assessment Update
January-February 1991
Volume 3, Number 1

assessment, faculty and student questionnaires, focus groups, and individual interviews are some of the strategies under investigation. The institutions involved are Alverno College, California State University, Lehman College, the University of Connecticut, Cuesta Community College, George Mason University, Miami University, Millsaps College, the College of William and Mary, and the University of Wisconsin Medical School. I was so excited about the work being done in this area that I came home and spent the next two weeks writing a review article for an issue of the *Journal of General Education*. The article devotes a mere half-dozen paragraphs to the use of standardized tests and then goes on to describe the variety of locally developed approaches that faculty are beginning to use.

In the assessment cluster, we decided that *Assessment Update* should have a column designed to facilitate dissemination of information about FIPSE assessment projects. "With FIPSE Support . . ." is a fitting title, I think, and it is also fitting that the initial appearance of this feature should coincide with the inauguration of *Assessment Update* as a bimonthly publication. Our thanks to Jossey-Bass Publishers for giving us this opportunity to become more responsive to our readers' needs for current information.

There is a variety of locally developed approaches that faculty are beginning to use.

TQM Without Assessment? How?

It was apparent that interest in TQM in higher education was spreading like wildfire.

Periodically, my interest in the content of a particular issue of *Assessment Update* causes me to postpone writing in this column about whatever topic is uppermost in my mind and devote the space instead to some observations about what our contributors have to say. This time, however, I don't have to put off writing about the subject on my mind, because it is Total Quality Management. The lead article, by Peter Ewell, is about that very issue.

In 1988 John Harris, a member of the *Assessment Update* editorial board who is now at Samford University, convinced me that I should begin to study the work of Edwards Deming and others on the subject of improving quality. I became so interested that in 1989 Harris and I collaborated on a FIPSE proposal that would involve creating a consortium of institutions to study the feasibility of implementing quality improvement principles in higher education. In that proposal—which FIPSE funded for a three-year period ending August 31, 1992—we pledged to provide a "translation" of quality improvement principles from the language of business and industry to that of higher education. Further, we said we would try to implement some of these principles in several of the institutions within the consortium. We have in fact made progress on both fronts.

Using the language of the academy, Harris and David Sylwester, who is head of the Department of Statistics in the College of Business at the University of Tennessee, Knoxville (UTK), wrote a 24-page paper, "Quality Improvement in Higher Education: An Overview," which provides a brief introduction to quality improvement principles—now more popularly called Total Quality Management (TQM)—for college and university faculty and administrators. (The paper is available at cost, $2.50, from Assessment Services, University of Tennessee, 1819 Andy Holt Ave., Knoxville, TN 37996-4350.) Consortium representatives at Samford, UTK, and Austin Peay State University have initiated pilot programs aimed at applying TQM.

Assessment Update
September-October 1992
Volume 4, Number 5

Last winter, as I contemplated what a culminating conference for the FIPSE project should attempt to accomplish, it was apparent that interest in TQM in higher education was spreading like wildfire. Subsequently, TQM sessions at the spring and summer higher education conferences drew standing-room-only crowds. I feared that a widely advertised conference on the topic would attract an enormous number of people, many of whom would be frustrated because so few institutions have enough experience in implementing quality improvement principles to provide concrete examples for others to study.

Consequently, we opted for a very small invitational conference designed to (1) provide a critique of the Sylwester-Harris manuscript, (2) take stock of where we are in implementing TQM in higher education, and (3) suggest where a funding agency like FIPSE might focus its resources in order to encourage the use of quality improvement principles in colleges and universities. We invited representatives of six research universities that are at least two years into TQM programs to join representatives of four of our consortium institutions for a meeting in Knoxville on May 22, 1992. Ewell's article summarizes the most important matters discussed at that conference.

From the time of my first encounter with the quality improvement literature, I have been convinced that on many campuses outcomes assessment would attain its highest value when viewed as integral to the data-gathering strategies that are so fundamental to TQM. I have long maintained that assessment cannot endure on a campus as an activity conducted for its own purposes alone but must be linked with ongoing practices such as institutional planning, curricular review, student development programming, and comprehensive peer review, which are considered vital to the life of the institution. Now we may add TQM to this list of vital activities with which assessment can and should be connected.

Assessment cannot endure on a campus as an activity conducted for its own purposes alone.

Interestingly, most of the research universities that have taken an early lead in implementing TQM do not have strong student outcomes assessment programs (see my column in the Vol. 4, No. 4 issue of *Assessment Update*, p. 3). Perhaps this, along with the fact that by and large the campus coordinators for our select set of TQM programs have backgrounds in business-related disciplines, makes it less likely (as Ewell found to be the case) that these TQM leaders will readily see ways to link assessment and TQM and apply TQM in academic, as well as administrative, campus functions. To their credit, they are eager to support TQM applications in curriculum development and instruction, but they do not see ways to do that immediately.

I would like to suggest that those of us who do know assessment and can see how a systematic approach like TQM can be strengthened by the data-gathering resources we already have in place, should begin to take an active part in the discussions of TQM on our own campuses. The idea of the campus itself as a learning system—a living, breathing *example* of the rational approach to seeking truth, which is ostensibly the basis for work in our disciplines—a powerful one. Some campus proponents of TQM are finding that group discussions of *The Fifth Discipline—The Art and Practice of the Learning Organization*, by Peter M. Senge of MIT

Assessment Update
September-October 1992
Volume 4, Number 5

(Doubleday, 1990), are helpful in understanding the implications of the university as a learning system.

In making our voices heard in the TQM dialogue, we can also draw on an increasing number of examples of applying quality improvement (and assessment) principles in curriculum design and classroom instruction. I recently attended a national conference of statistics professors in which examples from colleges of business at Minnesota, the Rochester Institute of Technology, and the University of Chicago were discussed.

More and more often I hear colleagues in assessment express concern that campus TQM initiatives are going to absorb funds that might have been allocated to support assessment. My response is to ask, "But how can you implement TQM without a strong assessment component?" I hope many of you are asking the same question.

How can you implement TQM without a strong assessment component?

Assessment Update
September-October 1992
Volume 4, Number 5

TRUDY W. BANTA

A Final Report to FIPSE on Assessment and Total Quality Management

Faculty and administrators considering outcomes assessment are concerned about the quality of methods available for gathering evidence of student learning and about the difficulty of using such evidence to improve the environment for learning. Most look within their own institutions and the higher education community in general for appropriate methods for collecting data and utilizing information. At the University of Tennessee, Knoxville, we decided in 1989 to seek FIPSE support to bring together the outcomes assessment coordinators from seven diverse institutions for the purposes of (1) improving their methods for gathering data from alumni and employers, (2) increasing their use of assessment information in making institutional improvements, and (3) studying the implications of total quality management (TQM) concepts utilized in business and industry for further enhancing the quality and use of assessment information in a process of continuous improvement within postsecondary institutions.

> *Ways must be found to link assessment with ongoing processes that are valued in the academy.*

Skepticism abounds among faculty concerning both the reason for doing outcomes assessment and its ultimate usefulness on campus. Ways must be found to link assessment with ongoing processes that are valued in the academy. The principal reason for undertaking the FIPSE project Applying Deming's Quality Improvement Strategies to Assessment in Higher Education was to explore the possibility that outcomes assessment could be accorded more status within the framework of institutional decision-making if it were associated with TQM.

While participating in the study of TQM, assessment coordinators at the State Technical Institute at Memphis, Roane State Community College (Harriman, Tennessee), Austin Peay State University (Clarksville, Tennessee), University of Tennessee, Martin, Memphis State University, University of Tennessee, Knoxville (UTK), and Samford University (Birmingham, Alabama) collaborated to improve the quality and relevance of an alumni survey that all public Tennessee institutions were required to use for the state's Performance Funding Program. These coordinators also developed an employer survey containing some of the same questions as those asked

Assessment Update
March-April 1993
Volume 5, Number 2

BEST COPY AVAILABLE

of alumni and sent the new survey to employers of alumni who gave permission for the employer contact as part of their response to the alumni survey.

Background and Origins of the Project

Tennessee's requirement that all institutions engage in prescribed outcomes assessment activities and report results as part of the state funding process provided a fertile context for this project. Moreover, the Center for Assessment Research and Development at UTK was able to contribute research expertise and outlets for dissemination of project findings.

John Harris, who moved just before the project began from a Tennessee institution to Samford University, was encouraged by the president and other leaders at Samford to launch a comprehensive TQM program on that campus. The laboratory for experimentation thus created at Samford provided a rich learning environment for the other institutional project participants.

On each field test, the mailed employer survey elicited a response rate in excess of 80%.

Project Description and Results

Over the three years of this FIPSE project, which ended in August 1992, two alumni questionnaires, two questionnaires for employers, and three methods of administering these linked surveys were tested. We found that the most effective and efficient methodology involved using a mailed alumni questionnaire accompanied by a form granting permission for employer contact, followed by a mailed survey form for employers that contained a forced-choice response format. On each field test, the mailed employer survey elicited a response rate in excess of 80%.

John Harris and David Sylwester, head of the Department of Statistics at UTK and TQM consultant to business and industry, jointly explored the application of TQM within higher education and developed a document that translates TQM concepts from the language of industry to that of academe. All participating institutions initiated TQM projects of one kind or another during the grant period.

Summary and Conclusion

Distances and politics in Tennessee had kept assessment coordinators from coming together to pool ideas and strategies prior to the inception of this project. By the time the project ended, relationships among the assessment coordinators were so firmly established and the benefits of collaboration so clearly demonstrated that at the last consortium meeting plans were made to form a statewide organization to continue the FIPSE-initiated tradition.

Assessment Update
March-April 1993
Volume 5, Number 2

TRUDY W. BANTA

Whereas on some campuses across the country 1992 brought fears that TQM would rob outcomes assessment of attention and resources, at the seven institutions involved in this project assessment coordinators provided leadership for both initiatives. Careful scholarly work on two data-gathering strategies has enhanced the credibility of at least some of the assessment methods being used by the participating institutions, and many more programs designed to improve the learning environment for students on these campuses are now based on assessment findings than was the case when the project began. This study demonstrated that given the right context, TQM can in fact increase the use of assessment findings in a process of continuous improvement in colleges and universities.

TQM can in fact increase the use of assessment findings in a process of continuous improvement in colleges and universities.

Assessment Update
March-April 1993
Volume 5, Number 2

An Audience of Beginners Needs Basics

We must be sure that our literature—especially Assessment Update*—contains plenty of basic information about assessment practice that will be helpful to the newest members of the audience.*

The *Assessment Update* editorial board meets annually in conjunction with the assessment conference sponsored by the American Association for Higher Education (AAHE). During these meetings we often reflect on the state of the field of assessment. When we convened in Washington in mid-June for our 1994 meeting, we began this discussion with the observation that each year at least one-half of the 1,200 to 1,500 participants in the AAHE conference are new to assessment. While the number of individuals who might be considered assessment specialists continues to grow slowly, actions of federal and state governments designed to focus institutional attention on reporting and using outcomes data annually cause increases in the number of colleges and universities and the number of individual faculty considering a serious approach to assessment.

With the audience for assessment continually expanding, we must be sure that our literature—especially *Assessment Update*—contains plenty of basic information about assessment practice that will be helpful to the newest members of the audience. I was pleased when I looked over the contents of this issue and found much that addresses the needs of newcomers. That is not to say, however, that this collection of articles is not of interest to our readers who are experienced assessment practitioners. Seasoned assessors simply read with a different perspective: we are seeking nuances in others' experiences that can help us improve our own practice.

Rosemary Aten's article provides a useful description of the steps that might be taken in launching a campus assessment program. A first step in any assessment effort is to establish a purpose, and at Western Illinois University, assessment was designed to improve student learning. Visits with department chairs and faculty committees by the assessment coordinator ensured that this purpose was communicated throughout the institution, and groups of faculty and students were encouraged to collaborate in constructing assessment strategies for individual departments. The process of involving as many players as possible in planning and implementing assessment is vital.

Assessment Update
September-October 1994
Volume 6, Number 5

BEST COPY AVAILABLE

but since policy and management structures differ from campus to campus, no single set of recommendations for proceeding will work in every setting.

Aten assembled a resource manual to provide clear direction for department faculty who were writing their own assessment plans. Many institutions have developed such suggestions for getting started and are quite willing to share these with others who are beginning their work in assessment.

Setting educational goals with specific learning outcomes for each is a first step in assessment at the department level. Then comes the crucial process of selecting or developing appropriate assessment methods for each outcome. Janet E. Boyle describes the portfolio, one of the most comprehensive and increasingly popular methods for assessing student learning. Her article contains the very important observation that assessment methods are most effective where they are structured as an integral part of other vital ongoing processes such as classroom instruction and testing.

Linda B. Rudolph, Daniel J. Poje and Janice Van Dyke offer advice on selecting assessment methods. They also summarize advantages and disadvantages of standardized versus locally developed measures. They point out that the most useful assessment data are derived from multiple measurement methods as opposed to a single technique.

Setting educational goals with specific learning outcomes for each is a first step in assessment at the department level.

Assessment data do not speak for themselves. The data must be analyzed, interpreted, and reported. Rudolph, Poje, and Van Dyke also enumerate some of the considerations that apply in deciding how to prepare and disseminate assessment findings. These authors also characterize the difficulties that may be encountered in encouraging faculty and administrators to follow through and use the results of assessment to improve teaching and learning and campus services. Their descriptions of the variety of uses to which assessment data may be applied are particularly helpful.

Alan N. Crist reminds us of the importance of focusing on educational processes as well as outcomes. Outcomes data such as test scores do not in and of themselves tell us much about what to do to improve student achievment. We must also look at methods of instruction, student motivation, time spent studying the material, and other intervening processes in order to obtain clues about what to try to fix when student achievement disappoints us.

Finally, assessment programs should be evaluated continuously and improved in accordance with suggestions provided by the resulting data. Tricia McClam and Thomas L. Bell describe an assessment of peer review that revealed that faculty consider the rewards for good teaching "grossly inadequate." In response, the campus chancellor established three new teaching awards and a rank of teaching scholar to match the rank of research scholar that is conferred annually on two extraordinarily gifted members of the faculty.

While framed in terms of the basics that newcomers to assessment need to know as they inaugurate their own programs, the foregoing paragraphs

Assessment Update
September-October 1994
Volume 6, Number 3

actually summarize many of the principles of successful practice in higher education assessment that we have derived from a decade of experience in this new field. We even have our own "Principles of Good Practice for Assessing Student Learning," developed in 1992 by a dozen assessment practitioners and scholars (see *Assessment Update*, 5(1), 7). But I sense that we need a larger work that consolidates as much as possible of what we know about effective assessment practice. I hope to develop such a compendium, complete with an abundance of examples to illustrate each principle that we can now enunciate.

I welcome readers' assistance in assembling the examples for this volume, which is tentatively titled *Assessment Strategies That Work*. If you have an approach that has worked at your institution, please send me a brief description (one to five double-spaced pages) organized according to the following outline: setting (description of your institution), purpose or purposes for assessment, description of the method or strategy, and an indication of how the assessment findings have been used. Contributions that arrive after the book goes to press will be considered for publication in *Assessment Update*. The new book and this periodical should help to ensure that our growing audience of beginners has access to the basic tools needed to construct an assessment program de novo.

I sense that we need a larger work that consolidates as much as possible of what we know about effective assessment practice.

Assessment Update
September-October 1994
Volume 6, Number 5

TRUDY W. BANTA

Revealing the Results of Assessment

As campus assessment coordinator at the University of Tennessee, Knoxville (UTK), I spent a great deal of my time communicating the results of assessment to various campus audiences. I believe that this activity was among the most important of my responsibilities because even the most significant findings may go virtually unnoticed if no one points them out to those who could use them.

We began each academic year by distributing to the academic affairs staff and deans a volume we call the *Detailed Summary of Instructional Evaluation Data*. It includes a one- or two-page synopsis of the findings of each of the major data-gathering activities we conducted as part of the assessment program during the previous year. The 1990 volume contained the results of testing seniors in general education and in selected major fields, of reviewing several master's degree programs, and of surveying samples of freshmen, enrolled undergraduates, undergraduate dropouts, seniors, and alumni.

In October 1990, we inaugurated an assessment column in the biweekly newspaper distributed to all faculty and staff. This provided an opportunity for broader distribution of some of the findings in the *Detailed Summary*.

Over the years, we tried to tailor the content of our data-collection strategies to reflect many of the broad goals contained in college- and university-wide planning documents. Annually, we developed a report for the deans and the Planning Committee that provided evidence of progress in meeting the relevant goals. For instance, as a research university, we worried so much about students' access to faculty that we made "to increase faculty-student interaction" one of our institutional goals. As evidence of progress in this direction, in 1983 a survey of undergraduates revealed that 49% knew no faculty members well enough to ask for recommendations; by 1990, this percentage had fallen to 27%.

Periodically, faculty and/or staff committees are appointed to study particular issues, such as student retention, the experience of older students, women's concerns, or campus library resources. With our survey responses, we recorded sufficient background information to be able to fashion special reports for these groups. One finding was of particular interest to the

I spent a great deal of my time communicating the results of assessment to various campus audiences.

Assessment Update
May-June 1991
Volume 3, Number 3

BEST COPY AVAILABLE

library committee: as a result of the opening of a new facility in 1988, students' library use—occasional or frequent—rose, from 53% in 1983 to 90% by 1990. These special reports also proved helpful to departments, colleges, and the university in providing concrete evidence for accrediting agencies of effectiveness in meeting specified goals.

By far the most effective channel for communicating assessment results at UTK was the comprehensive program-review process. Peer review is taken very seriously at UTK. One year before a visit by peers, faculty in a department begin to develop a self-study that outlines goals, curricula, and perceived strengths and weaknesses. Two outside specialists in the discipline are joined by three UTK reviewers outside the department for a visit of two and a half days. The visit includes interviews with virtually everyone associated with the program, from the chancellor and his staff to faculty and students and even heads of related departments. Reviewers provide written recommendations, and departmental faculty consider these and develop a written response that becomes the object of consideration at a follow-up meeting, which involves the responsible dean and appropriate representatives of the central administration. All concerned officials eventually sign an agreement specifying how each one will help the department carry out the actions necessary to address the recommendations considered most important. In a recent, almost extraordinary, attempt to ensure follow-up action, UTK has instituted a midcycle review. Full-scale reviews take place every eight years, but during the fourth year between these reviews, one of the outside reviewers is brought back to conduct, along with the three UTK reviewers, a visit of a day and a half, which once again brings the attention of the chancellor and his staff, the dean, and the department to the business of responding to the earlier recommendations for strengthening the department.

The most effective channel for communicating assessment results at UTK was the comprehensive program-review process.

In 1983, we added to the guidelines for the departmental self-study the section "Evidence of Program Effectiveness." Prior to 1983, the guidelines had asked the department to describe its resources and processes for achieving goals, but not its goal-related outcomes.

Also in 1983, my office began to collect assessment evidence for each department, evidence that could be included in its self-study—general education and disciplinary testing results, with appropriate comparative data, and targeted survey responses obtained from students, faculty, and alumni. We provided detailed reports for perusal by departmental faculty as well as brief summaries that can be included without modification, if desired, in the self-study document. This information can enhance faculty's and reviewers' understanding of the extent to which the department is achieving its stated goals.

Revealing the results of assessment to faculty at the time of program review proved to be the mode of communication most likely to produce lasting change at UTK. Departments gave increased attention to internships, became more active in placing graduates, revamped introductory courses, instituted departmental newsletters, and even changed curricula as assessment data focused the attention of faculty and reviewers on areas in need of improvement. This internal use of data in the formative evaluation of academic units is perhaps the most significant contribution that assessment has to make in higher education.

Assessment Update
May-June 1991
Volume 3, Number 3

TRUDY W. BANTA

The Power of a Matrix

Recent experiences at Iowa State University and on my own campus have convinced me to write about a tentative conclusion I reached about four years ago. I have come to believe that a simple matrix can be a powerful tool for helping faculty plan and organize an approach to assessment. Even a concept as abstract and complex as assessment can be made more concrete and comprehensible as relationships among its component parts are depicted visually in a single graphic image.

A faculty committee with responsibility for assessment at Iowa State used a matrix to enable department faculty to conceptualize assessment of student learning in the major. At Indiana University-Purdue University Indianapolis (IUPUI), general education is the current focus of a series of matrices developed to guide our work. Drawing on these experiences, I construct examples here to illustrate how the use of matrices can assist in planning for assessment with respect to both the major and general education.

The matrix is a versatile construction. It can be very simple or quite complex, ranging from a single element tracked over two dimensions (a 1-by-2 matrix), to two elements with two dimensions each (a 2-by-2 matrix), to many elements with many dimensions (X-by-Y matrix).

An excellent way to begin any faculty discussion of assessment is to ask what students should know and be able to do when they complete a course of study. Such a conversation might produce a list such as the following: (1) write effectively, (2) speak persuasively, (3) manage time, (4) use a computer to acquire new information in the discipline, and (5) demonstrate the ability to apply theory in a practical context.

The next step might involve creating a 1-by-1 matrix for each of the five skills in which the skill is broken into a set of subskills. Table 1 contains an analysis of the time management skill as conceptualized by the faculty of the Indiana University School of Nursing. Within the cells of this matrix, the nursing faculty have indicated the courses or other experiences in which

I have come to believe that a simple matrix can be a powerful tool for helping faculty plan and organize an approach to assessment.

Assessment Update
July-August 1996
Volume 8, Number 4

BEST COPY AVAILABLE

students will be expected to develop the subskills they have identified. A variation on the skill-by-subskills matrix is the skill-by-courses matrix in Table 2.

Table 1. Skill-by-Subskills Matrix

Skill	*Subskills*		
Time Management	Manages own time	Manages one patient's care in specified time	Manages care of several patients in specified time
	Courses 101, 120	Clinical experience	Hospital setting

Table 2. Skill-by-Courses Matrix

Skill	*Courses*		
Time Management	Course 101	Course 201	Course 401
	Manages own time	Manages one patient's care in specified time	Manages care of several patients in specified time

A matrix could be developed to illustrate the critical skills and subskills all students are expected to develop by the time of graduation.

Either of the simple matrices just described could be made more comprehensive by adding dimensions. Continuing with the five-skills example, a 5-by-X matrix could be developed to illustrate the critical skills and subskills all students are expected to develop by the time of graduation, along with the courses and other experiences that are designed to help students learn these skills.

Next comes the measurement question: What evidence do we need to convince us that students have learned the skills and knowledge we deem critical? Here a multidimensional subskills-by-measures matrix such as Table 3 can be helpful.

Table 3. Subskills-by-Measure Matrix

Subskills	*Measures*				
	Student Self-Assessment	Faculty Rating Scale	Supervisor Observation	Written Exam	Patient Questionnaire
Manages own time	✓	✓			
Manages one patient's care in specified time	✓		✓		
Manages care of several patients in specified time	✓		✓	✓	✓

Assessment Update
July-August 1996
Volume 8, Number 4

The rows of Table 3 assure us that we have multiple measures of each of the subskills; the credibility of our evidence is enhanced if it comes from several sources. The columns of Table 3 indicate the specific kinds of information each measure must contain if it is to fulfill its role in the assessment plan. For instance, in a self-assessment instrument designed for student use, we should include items that ask students if they feel confident that they can manage their own time, care for a single patient, and care for several patients effectively; then they might be asked if their program of study has contributed much or little to their expressed levels of confidence.

Most faculty want to know how they will find the extra time they believe will be needed to do the kind of comprehensive assessment of student learning that external demands for accountability require. The use of matrices can help faculty see that they are already collecting in their own courses much of the evidence needed for credible, comprehensive assessment. Moreover, the faculty development that occurs during conversations about assessment prompted by matrix construction can lead to improvements in course and curriculum structure and in pedagogy that make faculty work not only more effectively but also more efficiently, thus saving time in the long run. Combining the information in a skills-by-courses matrix with that in a subskills-by-measures matrix (see Tables 2 and 3) can show not only where skills are to be taught but also where they can be assessed. Periodically, a faculty committee overseeing the comprehensive assessment of students' knowledge and skills can look at samples of student work that have already been evaluated for purposes of giving grades and reassess these samples to derive a general sense of whether most students are developing given skills at the level or levels that faculty find acceptable for program graduates. (Incidentally, the interest in looking across students and across courses to assess student development in a generic sense suggests the need for yet another type of matrix: the skills-by-standards for judgment matrix in which excellent, good, fair, and poor [for example] levels of performance for each skill are carefully described.)

The use of matrices can help faculty see that they are already collecting in their own courses much of the evidence needed for credible, comprehensive assessment.

The skills-by-process matrix developed by faculty at Iowa State University is a very effective tool for tracking progress in assessment at department and campus levels. In column 1, a department lists its desired learning outcomes for students; in column 2, the courses and other experiences where the outcomes are taught are specified; column 3 contains the measures associated with each outcome; column 4 contains a summary of assessment findings; and, in column 5, faculty are asked what changes, if any, they have undertaken in response to the findings.

Looking at the five-column matrix constructed by each of Iowa State's departments, we can see many of the benefits that a conscientious approach to outcomes assessment confers. Bringing faculty together for collective consideration of desired student learning outcomes produces a better understanding by faculty and students of the role each course can and should play in overall student development. This understanding can create better focus in teaching, more deliberate sequencing of concepts, and thus improved learning. If students and graduates, employers, and other

Assessment Update
July-August 1996
Volume 8, Number 4

community representatives are involved in the goal-setting process, the experience can be even richer and the outcomes even more relevant, current, and motivating for students.

Consideration of tools and methods for assessment of comprehensive student outcomes can help faculty learn more about setting criteria for judgment and developing more valid assessment methods for use in their own courses. Once again, this activity can be enriched by the involvement of students, graduates, employers, and others. Students can tell us how questions should be worded to communicate the intended meaning. Those outside the academy can suggest assessment activities and contexts and even participate as assessors.

Finally, in addition to making constructive contributions to curriculum development, course sequencing, pedagogy, and assessment, matrices can help to improve student advising. The skill-by-subskill and skill-by-courses matrices can be used in orientation and subsequent advising sessions to help students visualize the curriculum and understand what they are expected to know and be able to do as they progress. With the addition of the skills-by-measures matrix, students learn when and how their achievement will be assessed. Having these blueprints should help students develop a better sense of purpose and thus become more highly motivated.

The use of a matrix series can produce improvements in current practice that continue to benefit students and faculty in a continuous cycle of improvements.

Iowa State's five-dimensional matrix focuses attention on the long-range and iterative nature of assessment: Measures yield findings that call for responsive actions, which, in turn, suggest revisions in desired learning outcomes. Thus, the use of a matrix series can produce improvements in current practice that continue to benefit students and faculty in a continuous cycle of improvements.

Assessment Update
July-August 1996
Volume 8, Number 4

Where Are We in Assessment? This Issue Says It All

TRUDY W. BANTA

I think this is one of the best *Assessment Update* issues ever. As one who filters everything through the lens of "How can I *use* this?," I often worry about whether a given issue contains sufficient information for practitioners on specific assessment techniques and how they are working. This issue is full of stories about campus experiences, and, together, they have much to tell us about the current status of outcomes assessment in higher education.

First, we might generalize from the content of this issue that faculty at two- and four-year institutions are using a wide variety of assessment techniques, ranging from commercially available survey instruments to faculty-developed written assignments, portfolios, and interviews. The ratio of faculty-developed instruments to commercial instruments as subjects of *Assessment Update* articles tells us something about the status of the field: the variety of faculty initiatives is growing faster than is the adoption of commercial methodologies. One of this issue's feature articles reminds us that assessment must be closely linked to important ongoing institutional processes if it is to realize its full potential in improving institutional effectiveness. Another article calls our attention to some of the continuing concerns of faculty about the ethical implications of our work in assessment. Both articles discuss topics upon which all of us who are deeply involved in outcomes assessment reflect from time to time.

I would like to say a little more about the features of this issue that stand out in my mind after reviewing them. It is significant that Gary Pike's column on the ACT/Evaluation Survey Service is the only piece in the issue that focuses upon a commercially available, standardized methodology. Many institutions begin their assessment programs with a commercially developed survey because it is readily available and can provide some useful data for consideration while faculty are developing their own measures. The ACT surveys are among the most popular instruments in this category. Another very helpful series of related surveys is the Program Self-Assessment Survey

The variety of faculty initiatives is growing faster than is the adoption of commercial methodologies.

Assessment Update
July-August 1991
Volume 3, Number 4

BEST COPY AVAILABLE

program offered by the Educational Testing Service, which includes instruments for students, faculty, and alumni and is designed to be used in the self-study phase of program review.

The faculty at Fairleigh Dickinson University (see the Memo, "Literature of Black Authors . . .") have developed a very simple yet powerful approach to assessment. It consists of asking students to identify three of the most influential ideas they have encountered in their core curriculum experience and to explain in a brief essay the intellectual significance of the idea and its personal impact. Then the students must rate the extent of personal impact on a five-point scale. What a rich set of data this approach can yield for gauging the relative importance of students' educational experiences!

At Emporia State University (see the Memo, "Portfolios Measures . . ."), the faculty have adopted an experimental approach to the assessment of the general education program that includes standardized tests, alumni surveys, and locally developed exams. However, Lendley Black reports that portfolios seem to be the method of choice in providing "a comprehensive view of student learning over a period of time."

I am convinced that conversations with students are essential elements in deepening our understanding of the results of assessing student outcomes.

At J. Sargeant Reynolds Community College (see the Memo, "Focus Groups . . ."), faculty use focus groups to collect data for assessing the general education program. I am convinced that such conversations with students are essential elements in deepening our understanding of the results of assessing student outcomes, regardless of the techniques we use.

Two of the articles in this issue illustrate how students can become active participants in assessment. At Indiana University-Bloomington (see the Memo, "Faculty Conduct . . ."), undergraduates have been trained to interview fellow students to determine how they allocate their time among studies and other activities. Carl Waluconis describes a self-assessment class at Seattle Central Community College in which students first take an inventory that helps them become more aware of their own learning styles, then analyze their learning experiences in all their classes, both orally and in writing. This article is enriched by its use of students' own written comments in illustrating their development as self-assessors.

Robert J. Barak elaborates the important point that assessment can be made more meaningful, useful, and important if it is incorporated in a comprehensive academic program review process. Linda Rudolph and Daniel J. Poje call for increased attention to relating assessment methods to specific purposes, studying the reliability and validity of assessment instruments, protecting the privacy of students participating in assessment, and exercising care in analyzing and interpreting assessment data. This message from my fellow Tennesseans was particularly poignant as we engaged in dialogue with our state coordinating board about the use of differences as small as one-tenth of a point in test scores or alumni survey responses in making performance funding decisions that determined the gain or loss of thousands of dollars of state funding for each public institution in the state.

Assessment Update
July-August 1991
Volume 3, Number 4

TRUDY W. BANTA

In a follow-up to the 1990 ACE *Campus Trends* survey, ACE and Winthrop College staff discovered that only 16% of campus assessment coordinators have a degree in an assessment-related field, and 76% say they rely on assessment workshops and seminars for their training. National assessment leaders now believe that training is the field's preeminent need.

Only 16% of campus assessment coordinators have a degree in an assessment-related field, and 76% say they rely on assessment workshops and seminars for their training.

Assessment Update
July-August 1991
Volume 3, Number 4

Terenzini's Purposes Frame This Issue

The power of accountability is a driving force behind much of the activity in assessment.

At the Fourth International Conference on Assessing Quality in Higher Education held in 1992 in Enschede, Netherlands, Jim Ratcliff and Pat Terenzini, co-directors of the National Center on Teaching, Learning, and Assessment at The Pennsylvania State University, were invited to present the U.S. perspectives on assessment. (Peter Ewell and Dennis Jones of the National Center for Higher Education Management Systems had this responsibility at the Fifth International Conference held in Bonn, Germany, July 19-21, 1993.) In lieu of preparing a text in advance of the meeting, I suggested that Pat develop for the closing session a summary based on his observations during the conference. As the meeting got underway and I began to appreciate the diversity of worldwide developments in assessment that had occurred during the year since the Third Conference in England, I wondered if I had proposed an impossible task for my friend from Penn State. How could he draw together and make sense of so many disparate approaches in the space of two days?

As is characteristic of all Pat Terenzini's work, his summary was excellent. And he delivered it with confidence, grace, and wit, without once chiding the conference organizer from the United States for having spoiled his enjoyment of an all-too-brief visit to the Netherlands with such a demanding assignment. We published the full text of Pat's remarks in the conference *Proceedings*, and I am pleased to include an edited version as the feature article of this issue of *Assessment Update*.

Terenzini's paper provides an excellent context for the subsequent material in this issue. He affirms the importance of philosophy and purpose for assessment, characterizing accountability and improvement as the two principal purposes. Sal Corrallo, in his article on the national assessment of college student learning, reminds us again of the power of accountability as a driving force behind much of the activity in assessment. Corrallo tells us that the National Center on Education Statistics is continuing its efforts to develop a working consensus regarding the workplace and citizenship skills needed by today's workers and citizens and that this is a preliminary

Assessment Update
May-June 1993
Volume 5, Number 3

BEST COPY AVAILABLE

activity in the development of a national assessment of the attainment of these skills.

Peter Ewell's "From the States" carries the accountability theme from the national to the state level. He reports that recently "legislation by fax" has produced similar bills in several states that compel institutions to begin reporting data on performance indicators.

Assessment for the purpose of improving the educational experience for college students is the focus of virtually every other article in this issue. In their presentation of the Hampton model, Linda Petty, Eleanor Lynch, and John Alewynse describe a course X objectives matrix model to connect institutional, departmental, and course objectives. In my own work with faculty groups, I have found the matrix concept to be a very helpful learning device as we explore ways to accomplish that all important step in assessment: ensuring that the objectives faculty think are important are actually taught.

In "Facilitating Empowerment Through Student Guidance," Nancy Gadbow discusses the importance of self-directed learning and assessment. She emphasizes the need for instructors to discover their own unique learning styles and recognize their influence on teaching so that students' learning styles may be addressed more self-consciously in classroom instruction.

I believe that the real value of assessment is coming to be its power to focus the attention of faculty on new ideas for improving the experience they provide for students.

Even the "Memos" and "Resources" sections of this issue include tips on improving teaching and learning. Susie Jans-Thomas from Mt. Mary College reports on a method for obtaining immediate feedback from students on the effectiveness of classroom lectures and other instructional strategies. Thomas Donlon reports on a student response program at Thomas Edison College that provides an avenue for criticism of the evaluation process at the time students encounter it.

Perhaps the best known work in higher education on the topic of using assessment to improve teaching and learning in the individual classroom has come from Tom Angelo and Pat Cross. As noted in "Resources," we are currently being treated to a second edition of the best-selling title by these authors: *Classroom Assessment Techniques: A Handbook for College Teachers*.

Finally, Kathy Baugher, formerly at Samford University and now at Belmont University, has prepared an exciting new publication, LEARN. This manual for students helps them develop teams that can provide valuable feedback to their instructors concerning the effectiveness of the processes used to promote learning in specific courses.

Over the last several months I have been most encouraged by the fact that the articles being submitted to *Assessment Update* increasingly concern use of the assessment process and results to improve teaching and learning. I believe that the real value of assessment is coming to be its power to focus the attention of faculty not heretofore well acquainted with the literature on effective pedagogy with new (for them) ideas for improving the experience

Assessment Update
May-June 1993
Volume 5, Number 3

they provide for students. While many faculty have had little training in the principles of good teaching, assessment is helping them discover these principles on their own—a powerful way to learn. When these self-discovered principles are confirmed in encounters with the work of higher education scholars that occurs in faculty development workshops, at national conferences, or in independent study of the literature, the learning becomes even more profound and lasting. Thus, as many of us now know from our own experience, engagement in assessment can be a very effective means of promoting faculty development.

While many faculty have had little training in the principles of good teaching, assessment is helping them discover these principles on their own—a powerful way to learn.

Assessment Update
May-June 1993
Volume 5, Number 3

TRUDY W. BANTA

This One's for Students

With great pleasure we present an issue in which the spotlight is squarely on students. Since our readers as well as our contributors are principally college and university faculty and administrators, *Assessment Update* is always filled with articles designed to help faculty understand and improve assessment practice. In this regard, this issue is no exception; the articles will enrich faculty perspectives. The difference is that instead of the usual focus of faculty writing about and learning from the experience of faculty colleagues, most of the articles herein describe ways in which faculty can learn by tapping the experience of students.

Faculty can learn by tapping the experience of students.

Nearing the midpoint of our seventh year of editing *Assessment Update*, I can look back and find only one previous issue in which the cover article featured students as key players in assessment: Douglas Lee Hall of St. Mary's University in San Antonio gave us the wonderful story "Involving Students as Active Participants in Assessment" (*Assessment Update*, 1993, 5 (5), 1-2, 6). As the new, overworked chair of computer sciences, Hall gratefully accepted an offer of help from his college work-study student. He asked her to assist him in thinking through the curriculum for a new bachelor of science degree that the faculty intended to propose. This experience was so positive that Hall asked his colleagues to nominate representatives for a computer science advisory board, to be made up solely of students. Using Association for Computing Machinery guidelines, the students developed a proposal for the new undergraduate curriculum that included rationales for all of the courses as well as for their sequencing. The department's faculty accepted the proposal with little change, and the university's academic council approved it with comments about "how well the curriculum had been thought out, how rigorous it was, and how reflective it was of the purpose and mission of the university" (Hall, 1993, p. 1).

The majority of articles in this issue of *Assessment Update* attest the value for faculty of involving students in assessment. The lead article is the first in a series of four reports on the use of Total Quality Management (TQM), or Continuous Quality Improvement (CQI), in higher education. Judy

Assessment Update
May-June 1995
Volume 7, Number 3

BEST COPY AVAILABLE

Griffith, John McLure, and Jann Weitzel take on one of the most vexing challenges faculty—the large class—and show how they have used TQM strategies to improve the student experience in a required teacher education course at the University of Iowa. For instance, a color-coded seating chart that includes some information about each student has helped make learning more personal, discussions more interesting, and classroom participation more universal. A simple knowledge checklist asks students to rate on a three-point scale their knowledge of important concepts that the course is designed to cover. The results enable the instructors to assign meaningful readings and to skim over certain concepts in order to spend more time on others. Assessment of written work with detailed marginal comments helps students improve their grades on papers and midterm exams if they are willing to revise their originals and resubmit them.

Speaking with the authority of a professor in the School of Business at the University of Indianapolis, L. Leslie Gardner observes in the second article that "efficiency is increased by eliminating the need for a separate time slot for a classroom assessment activity by combining that activity with the teaching of process improvement." Students spend some of their class time in groups of three or four, applying to the task of improving the class the CQI techniques of brainstorming, prioritizing ideas, constructing Pareto charts and fishbone diagrams, and conducting surveys. Students are graded on completion of a prescribed sequence of CQI activities and on the practicality of implementing the improvements they have suggested.

Relationships with peers can enhance the learning of individuals.

In team teaching a TQM course at Marietta College, Alice M. Warner and David Cress tossed aside the syllabus, text, and traditional methods of assessment and engaged in collaborative efforts with their students to design the course and appropriate assessment techniques. Both instructors and students developed portfolios that included individual and group learning goals, papers reflecting on processes and progress, class improvement feedback sheets completed at the end of each class, and concrete evidence that learning was taking place. An important aspect of the learning that occurred was public sharing of the personal evaluations as they were added to the portfolios. The instructors' reflection papers were particularly important to the students' understanding of how relationships with peers can enhance the learning of individuals.

The article by Paula M. Rooney and P. Gerard Shaw illustrates the increasingly popular CQI practice of benchmarking. When registration and the billing process topped the list of campus irritants in student surveys conducted at Babson College in 1989 and again in 1993, the Babson student affairs staff began looking for a college or university that had solved its registration problems. When they could not find good models in higher education, Babson staff turned to their food service provider, the Marriott Corporation. They found that Marriott's hotel affiliate had been working for years to improve its registration process, and that Marriott could provide benchmarks for Babson in this area. Using technology, Babson now handles most of the registration process by mail and telephone, thus eliminating waiting lines and "saving paper, time, and money." Continuous assessment

Assessment Update
May-June 1995
Volume 7, Number 3

of student opinion will enable Babson to tell if the improvements are satisfying its customers.

Joe V. Chandler and Bettie Horne report on the use of a particularly comprehensive set of questions in exit interviews for students in the Division of Physical Education and Exercise Studies at Lander University. Not only do the interviews solicit opinions about campus programs and services but they also ask about the strengths and weaknesses of the university's general education program and about the teaching effectiveness of individual professors in the division. At the end of the interview, the questioner—always someone from outside the division—rates the student's verbal communication skills. Students' complaints about the difficulty of scheduling certain courses convinced faculty to teach most courses annually as opposed to every other year, two courses thought to contain needless overlap were revised, and some negative evaluations shared privately by the division chair with a single professor led that individual to make changes in teaching methods.

Dennis E. Peacock gives us an early report on a campuswide student advisory group at Northeast Missouri State University that helps "assess assessment." In its first year this body has worked to improve (1) the design, content, and return rate of student surveys; (2) the interpretation of survey findings and general education test scores; and (3) students' motivation to do their best work on the general education exams.

In a relatively new field like outcomes assessment, it is important for faculty to talk first among themselves about useful materials and techniques. The majority of articles here argue strongly for involving students in the conversation as soon as possible. Faculty know most about course content—what students should learn—but students know most about how they learn. They can tell us much about the teaching strategies that help them grasp the important content most efficiently and effectively. We can improve our practice by listening to them.

Faculty know most about course content—what students should learn—but students know most about how they learn.

Assessment Update
May-June 1995
Volume 7, Number 3

The Responsive Community College

Most community colleges are very responsive to the educational needs of their students and their communities.

We hear much these days about higher education being out of touch with its constituents, about faculty who care more about pursuing their own narrow research specialties than about listening to the educational needs of their students, employers, and society. While there may be some truth to these comments about faculty on four-year campuses, there is little reason for criticism of community college faculty in this regard. Most community colleges are very responsive to the educational needs of their students and their communities.

Kay M. McClenney, vice president of the Education Commission of the States and one of the first *Assessment Update* editorial board members, recently made this point in the on-line archive of *On the Horizon*. "Workforce training, moving people from the welfare rolls to the tax rolls, involving employers, responding quickly to community needs—are at the top of political agendas. . . . Community colleges are well-equipped to help in the solution to these problems."

In the lead article for *Assessment Update* (Vol. 7, No. 5), Sara M. Morris describes a multiphase assessment and follow-up process that helps Asheville-Buncombe Technical Community College respond to the needs of employers in its service area. Every five years, the college brings together for focus group discussions the chief executive officers (CEOs) of the major industries in its community. These employers are asked to identify the knowledge and skills needed by entry-level employees as well as major skill deficiencies in the current work force that may suggest the need for retraining or continuing education. The focus groups supplement information about placement and satisfaction levels that the college gathers on a more frequent basis from surveys of its graduates and their employers. The college also relies on an advisory committee of practitioners for each curriculum area to suggest equipment updates and curriculum modifications.

An elaborate follow-up process that begins as soon as the focus groups are concluded invites further dialogue with the CEOs who participated. The

Assessment Update
September-October 1995
Volume 7, Number 5

activity culminates in a presentation of findings and recommendations for all faculty, students, and staff by the focus group chairs. If a new training program is suggested, the idea is tested with the faculty and the advisory committee associated with the curriculum most likely to be affected. No final decisions are made about adding a program until all focus group participants, as well as eight or ten other influential employers in the field being considered, have been interviewed and a comprehensive survey on the proposal has been administered to 100% of the group of employers in that sector of the local economy. It is difficult to suggest any improvements that would make this community college's process more responsive to the needs of the employers in its community!

As I worked with the manuscripts submitted by chapter authors for the book *Making A Difference: Outcomes of a Decade of Assessment in Higher Education* (Banta and Associates; Jossey-Bass, 1993), I was most favorably impressed by the responsiveness to assessment findings that were reported by faculty and administrators associated with two-year institutions. In their chapter on methods for demonstrating the effectiveness of community colleges, R. Dan Walleri and Jeffrey A. Seybert described assessment processes such as student tracking and program review that have provided guidance for improvement actions on two-year campuses.

It is difficult to suggest any improvements that would make this community college's process more responsive to the needs of the employers in its community!

Lane Community College in Oregon serves a community in which layoffs in timber, lumber, fishing, and agricultural industries have produced many unemployed workers with limited educational backgrounds, deficient academic skills, and low self-esteem. Lane's student tracking system, containing over 50 data elements from admissions, testing, and academic records, has been used to identify students most at risk of dropping out. Using the information that poor reading and writing skills and uncertainty about selecting an academic major are powerful predictors of problems at Lane, faculty and staff have developed new advising procedures with intrusive counselor intervention if needed, and a college success course. The course has been so successful that students who complete it are much more likely to finish credit-bearing courses than those who do not.

Eastern Iowa Community College District conducts a systematic review of each of its programs on a three-year cycle. In addition to a self-study prepared by faculty, the program review incorporates constituent surveys and a local labor market assessment. Thus, the health and viability of college programs are evaluated continuously, recommendations for improvement are derived, and informed decisions are made concerning the need to strengthen or terminate programs. As a result of program reviews in the Eastern Iowa system, computer-aided design has been added to the drafting program; the automotive technology program has been authorized to award a certificate after one semester, a diploma at the end of a year, and an associate's degree upon completion of the two-year program; and customized contract programs for local business and industry have been developed in computerized numerical controls, statistical process control, and Total Quality Management (TQM).

Assessment Update
September-October 1995
Volume 7, Number 5

And speaking of TQM, I must add that community colleges are prominent among the leaders in higher education in establishing indicators of their effectiveness, gathering benchmark data, and using findings to improve the satisfaction of students and other community constituents. Two good illustrations are in *Assessment in Practice: Putting Principles to Work on College Campuses* (Banta, Lund, Black, and Oblander; Jossey-Bass, 1996). First, Jo Ellen Cantrell of Spartanburg Technical College tells how 35 effectiveness indicators permit student development services staff to assess program effectiveness, track progress, and make improvements. One of the indicators is the number of errors made in student financial aid records as a percentage of the files audited each term. Errors are classified by type and frequency and relevant training is conducted to enable staff to improve their work. An initial error rate of 19% was thus reduced to just over 5% in a single year. Second, James L. Hudgins, Dorcas A. Kitchings, and Starnell K. Williams describe improvement efforts at Midlands Technical College in Columbia, South Carolina. Among six areas considered critical to Midlands' success are post-education satisfaction and success of students as well as economic development and community involvement. Employers' evaluations of the preparation of employees who attended Midlands constitute one of the sources of data used by faculty to determine and revise course content. Midlands administrators chart progress on 20 indicators and include some of the findings in an impressive annual community report designed to demonstrate the institution's accountability to the public.

Readers interested in more information on the responsive community college may consult Jeffrey A. Seybert's column, *Community College Strategies*, which is a regular feature of *Assessment Update*.

> *Community colleges are prominent among the leaders in higher education in establishing indicators of their effectiveness, gathering benchmark data, and using findings to improve the satisfaction of students and other community constituents.*

Assessment Update
September-October 1995
Volume 7, Number 5

TRUDY W. BANTA

The Rising Star of Assessment Scholarship

Donald A. Schön, professor emeritus of urban studies and planning at Massachusetts Institute of Technology, argued in a recent issue of *Change* for more acceptance in research universities of action research or reflection on practice. According to Schön (1995, p. 30), "We need to observe ourselves in the doing, reflect on what we observe, describe it, and reflect on our description." He believes that good practice itself can become the foundation for generating new knowledge if it is conceptualized and written about in a way that permits others to draw generalizations from it for their own practice.

Assessment is certainly a form of practice that should be the subject of action research.

Assessment is certainly a form of practice that should be the subject of action research. But, until recently, few higher education assessment practitioners have had the opportunity to reflect on their work and write about it in ways that would permit generalizations to be derived. As recently as 1994, Marcia Mentkowski and others were identifying reasons for the paucity of higher education assessment research (see *Assessment Update*, Vol. 6, No. 1). Now we have evidence to suggest that this situation is changing.

Each year I write to all of the individuals who made presentations at the annual American Association for Higher Education assessment conference and ask them to consider submitting articles to *Assessment Update*. I am always gratified by the response we receive to this request, but this year I am exceptionally pleased. The feature articles in *Assessment Update* (Vol. 8, No. 6) illustrate the reason for my delight. All summarize the results of carefully designed action research.

In the lead article, Trudy Bers, Diane Davis, and William Taylor describe a relatively new assessment method: syllabus analysis. In this area, as in transcript analysis, some of the best pioneering work has been done at community colleges, yet the method is applicable throughout higher education.

Bers and her colleagues were prompted to conduct their study as a means of finding out whether learning assignments were stimulating students to

Assessment Update
November-December 1996
Volume 8, Number 6

BEST COPY AVAILABLE

engage in activities like writing, speaking, and thinking critically, activities that faculty believe enhance student achievement in general education. The investigators decided that syllabi would provide an opportunity to look at what faculty are communicating to students about their expectations for student learning, as well as their standards for achievement.

The Bers, Davis, and Taylor article also provides a good example of how faculty and staff can collaborate to produce particularly strong work in assessment. Bers is senior director of institutional research and thus is in a position to design and carry out a study that includes appropriate checks on instrument reliability and a replicable action research design. The study findings have provided a stimulus for faculty development in designing and writing class syllabi. A follow-up survey has been designed to help the investigators learn more about what students value in syllabi. And the syllabus analysis itself will be repeated in spring 1997.

Faculty and staff can collaborate to produce particularly strong work in assessment.

The second article in the issue is the second contribution to *Assessment Update* by Frank J. Spicuzza, who heads the bachelor of science in social work program at the University of Tennessee, Knoxville. In Spicuzza's (1990) first *Assessment Update* article, he described a new approach to assessing student achievement in the social work major. The article represents a logical step forward in that it involves an evaluation of the performance in practice of one component of that new approach: the student portfolio.

The social work faculty at Tennessee have taken portfolio assessment seriously. They have developed the standards for preparing students to create the portfolio and have communicated those early in the program. Then they have standardized the criteria faculty use in grading the portfolios. In the study reported here, two methods of evaluating portfolio use were developed and applied.

The findings confirm some often-expressed benefits that faculty associate with portfolio use. However, the student data also provide some interesting contradictions. Most important from the reader's perspective, this study gives us new tools to help us reflect on our own practice in portfolio assessment.

In the third feature article, Elizabeth A. Jones, a research associate at the Pennsylvania State University, summarizes research that she has conducted for the National Center for Postsecondary Teaching, Learning, and Assessment over the past four years. She poses the question, "Is there consensus about college outcomes among faculty, employers, and policy-makers?" By and large, the answer—derived through an iterative survey process—is yes. But some of the areas of disagreement are noteworthy. For instance, employers feel that it is important for students to learn to present information visually—using tables and graphs—as well as in writing. Faculty, on the other hand, express more interest in helping students learn the writing process—developing, drafting, and revising ideas. While faculty value the skill of public speaking, employers see less need for public speaking than for students to develop relationship skills, that is, the ability to work in teams

Assessment Update
November-December 1996
Volume 8, Number 6

and facilitate communication among members of a small group. Faculty can be informed by such findings as they attach specific definitions and criteria for student achievement to often vaguely defined generic skills such as communication and critical thinking.

If these three articles are any indication, we are now embarking on a second decade of widespread experience in assessment that will see many more experiments and documentation of their effects, and much more of the reflective practice Schön describes, than ever before. Our field will be much richer as this fundamental scholarship begins to mature.

We are now embarking on a second decade of widespread experience in assessment.

References

Schön, D. A. "Knowing in Action: The New Scholarship Requires a New Epistemology." Change, 1995, 27 (6), 27-34.

Spicuzza, F. J. "Two-Part Examination at UTK." *Assessment Update*, 1990, 2 (4), 1, 5-7.

Assessment Update
November-December 1996
Volume 8, Number 6

Engineering Criteria 2000: Focus on Outcomes

The focus of engineering accreditation changed from ensuring that an institution is providing acceptable educational programs and processes to ensuring that the institution's students are attaining specified learning outcomes.

The Accreditation Board for Engineering and Technology (ABET) is changing the standards it uses to accredit engineering education programs. Its *Engineering Criteria 2000*, adopted in November 1996, changed the focus of engineering accreditation from ensuring that an institution is providing acceptable educational programs and processes to ensuring that the institution's students are attaining specified learning outcomes. Obviously, this changes everything.

All engineering programs must be ready to address the new *Criteria 2000* by the turn of the century. Those programs preparing for accreditation visits during 1998 and 1999 may elect to be reviewed under the current standards or according to *Engineering Criteria 2000* (R. C. Seagrave and M. Dayne Aldridge, personal communication, Mar. 7, 1997). ABET officials suggest that institutions assess their readiness to undergo accreditation under the new standards by comparing their practices to a series of levels of development. Institutions at the higher levels are encouraged to use the new criteria. Level 5, the highest, includes the following institutional characteristics: "Sound, systematic, highly integrated process and outcomes improvements are being made, enabled by quantitative feedback and by piloting innovative ideas and technologies. A constituency-centered approach is evident throughout. Sustained, world-class outcomes are apparent in all major areas. Results are clearly influenced by the approach." (Seagrave and Aldrich, personal communication, Mar. 7, 1997).

The Colorado School of Mines (CSM) is well on its way toward level 5. Spurred by Colorado's accountability legislation of the late 1980s, Olds and Pavelich (1996) have written that CSM has "agreed on its overall educational goals and has articulated them in a Profile of the CSM Graduate; these goals determine what materials are collected for the portfolios we maintain and evaluate for selected students. The data compiled from annual portfolio reviews help the faculty to identify strong and weak aspects of our programs and address the latter in a continuous improvement loop." In the lead article of this issue (Vol. 9, No. 3) of *Assessment Update*, Barbara M.

Assessment Update
May-June 1997
Volume 9, Number 3

BEST COPY AVAILABLE

Olds and Ronald L. Miller describe a pilot test of an innovative approach to teaching engineering students to apply the principles of total quality management (TQM). Instead of adding a course on TQM, CSM faculty have integrated TQM principles in a four-semester sequence of introductory courses designed to promote communication skills, critical thinking abilities, and teamwork. Exemplifying the spirit of continuous improvement, the faculty have built into their work an evaluation component that assesses performance and employs a questionnaire to measure students' perceptions and attitudes toward the worth of TQM principles and concepts in their project work. In comparison with students in classes where TQM modules were not used, those in the pilot sections produced superior project solutions and project reports and exhibited more understanding of customer focus, teamwork concepts, quality tools, and the use of data to effect improvements.

Nina W. Brown of Old Dominion University (ODU), in the second article, joins Joseph A. Shaeiwitz of West Virginia University (see *Assessment Update* 1996, Vol. 8, No. 4, pp. 4, 6) in advocating the use of student work on senior design projects in outcomes assessment. Brown describes criterion 8, "Outcomes Assessment Process," of the *Engineering Criteria 2000* and identifies the ten abilities that ABET calls the "attributes of an engineer" (Joint Task Force on Engineering Education Assessment, 1996). Brown, who is coordinator of assessment in the College of Engineering and Technology at ODU, created a 5-point checklist based on the abilities described in the set of attributes. That checklist is used by faculty in three engineering fields and four technology fields to assess the content of student portfolios based on their senior design projects. Outcomes assessment in engineering and technology at ODU has helped faculty identify needs for curriculum revisions, instructional improvement, and faculty development.

It is no surprise that engineering and technology faculty are searching for assessment methods to apply in their own settings.

With ABET's new emphasis on the assessment of student learning outcomes, it is no surprise that engineering and technology faculty are searching for assessment methods to apply in their own settings. In the previous issue of *Assessment Update*, we announced in the Calendar section an April 1997 engineering assessment conference at Rose-Hulman Institute of Technology in Terre Haute, Indiana. The conference, "Best Assessment Processes in Engineering Education," was cosponsored by ABET and the National Science Foundation (NSF) and featured some of the work of a consortium of seven institutions, including Rose-Hulman, Texas A&M University, Arizona State University, University of Alabama, Maricopa Community College District, Texas Women's University, and Texas A&M at Kingsville. The consortium has received a NSF grant for $14 million over five years for the purpose of helping engineering faculty develop approaches to assessment.

In the third article, Sam Cotton, a professor in the Department of Industry and Technology at Ball State University, and his colleagues Reza T. Ahmadi and Russell A. Esselborn, describe their assessment strategy to help faculty select simulation games for use in instruction. For illustrative purposes, the authors have applied their strategies to games that promote the competencies identified in the Secretary's Commission on Achieving Necessary Skills 2000 work force goals.

Assessment Update
May-June 1997
Volume 9, Number 3

While not specifically stated in ABET's ten "attributes of an engineer," critical thinking is a skill that underlies several of the attributes. Anne Hummer writes, in the fourth article, about faculty-developed assessment of critical thinking. Hummer provides a very helpful step-by-step explanation of the strategy—from arriving at consensus on the meaning of critical thinking within a given discipline to deciding how to assess specific critical thinking skills in a capstone paper.

Gary Pike's Assessment Measures column also addresses the need for appropriate measures of critical thinking. He recommends carefully developed surveys for use by upper-division students in self-assessment of such skills as defining and solving problems, thinking analytically and logically, and seeing relationships among ideas. He also mentions the College Student Experiences Questionnaire (CSEQ) as a source of information about critical thinking abilities. In this issue's Memos column, George Kuh announces the publication of revised CSEQ norms.

Finally, we are pleased to publish a second *Assessment Update* contribution by Patricia D. Murphy, this time with colleague Jeffrey Gerst, on the subject of assessing graduate student outcomes. In the first article on this topic, Murphy (1994) merely whetted our appetite for more by providing the briefest of overviews of her work. Now we are treated to a much more extensive description of the use of primary trait analysis by faculty at North Dakota State University in defining and assessing successful performance in comprehensive exam responses, research proposals, oral seminar presentations, theses, and oral defenses.

We hope that the assessment of graduate student outcomes will prove helpful to our colleagues in engineering and technology disciplines.

We hope that the assessment of graduate student outcomes, as well as the approaches to measuring critical thinking, selecting simulation games, using portfolios in senior design projects, and evaluating the introduction of TQM modules in lower-division engineering courses, will prove helpful to our colleagues in engineering and technology disciplines. We wish them well in addressing the new *Engineering Criteria 2000*!

References

Joint Task Force on Engineering Education Assessment. "A Framework for the Assessment of Engineering Education." ASEE Prism, May-June, 1996, 19-26.

Murphy, P. D. "Assessing Student Learning in Graduate Programs." *Assessment Update* 1994, 6 (6), 1-2.

Olds, B. M., and Pavelich, M. J. "A Portfolio-Based Assessment Program." In *Annual Conference Proceedings of the American Society for Engineering Education*, 1996.

Schaeiwitz, J. A. "Capstone Experiences: Are We Doing Assessment Without Realizing It?" *Assessment Update* 1996, 8 (4), 4, 6.

Assessment Update
May-June 1997
Volume 9, Number 3

TRUDY W. BANTA

Of Websites and Portfolios

Technology: Best New Hope for Assessment?

The theme of the 1996 Assessment Conference in Indianapolis was "Technology: Best New Hope for Assessment?" Given the questions asked by participants, we should have included in the conference materials a guide to assessment information on the Internet. We had looked for such a reference but could not find a comprehensive one. Now Mark Connolly and Jane Lambert have provided this important resource in our lead article.

Connolly and Lambert, doctoral students in higher education at Indiana University, have special interests in assessment and consult the Internet regularly in connection both with their coursework and with their apprenticeships with faculty. Lambert has an assistantship in my office and serves as an assistant editor of *Assessment Update*, while Connolly works with George Kuh in Bloomington on the College Student Experiences Questionnaire, a very helpful assessment resource that Gary Pike has described in *Assessment Measures* (*Assessment Update*, 1990, Vol. 2, No. 1 on page 294).

As Connolly and Lambert point out, the exponential growth of information on the World Wide Web renders any listing of Internet sources a little out of date even as it goes to press. However, the authors have made their work as enduring as possible by including a number of well-established reference sites that will be kept current for the foreseeable future. These include the sites maintained by the Educational Resources Information Center, the Fund for the Improvement of Postsecondary Education, the National Center for Higher Education Management Systems, the American Association for Higher Education, and the American College Personnel Association.

The next three articles in this issue, those contributed by W. Tracy Dillon; Charles W. Spurr, Michael J. Kiphart, and Betty Jo Miller; and Sue Ellen Shay, provide evidence of the substantial current interest in using portfolios in higher education assessment. It has been three years since we mentioned portfolios in this column (*Assessment Update*, 1994, Vol. 6, No. 5, pp. 3, 15). At that time, Janet E. Boyle (*Assessment Update*, 1994, Vol. 6, No. 5, pp. 10-11) contributed an excellent primer on portfolios, suggesting

Assessment Update
March-April 1997
Volume 9, Number 2

BEST COPY AVAILABLE

purposes, definitions, and methodological considerations for those contemplating their portfolios. Boyle (1994, p. 10) summed up the rationale for the burgeoning attention to portfolios as follows: "The portfolio, as an element of authentic assessment, has captured the interest of many instructors who want a more comprehensive way to assess their students' knowledge and skills, to have students actively participate in the evaluation process, and to simultaneously develop students' skills of reflective thinking. These latter features make portfolios an attractive alternative to traditional summative testing."

As the articles in this issue suggest, many faculty have been persuaded to try portfolios, and their extensive experience has carried us far beyond the primer stage. In fact, a rather stable set of generalizations about portfolios as higher education assessment tools has emerged. Collectively, the three articles included in *Assessment Update*, Vol. 9, No. 2 illustrate a number of these.

Portfolio use requires effective planning.

Good assessment of all kinds—and portfolios are no exception—begins with the identification by faculty of goals for student learning in courses, course sequences, and entire curricula. If such goals have not been made explicit previously, a study of what portfolios have to offer can encourage faculty to develop them. Since, as Shay points out, portfolios are most useful in portraying student development over time, particularly from the beginning of a course of study to the end, their use can be quite effective in stimulating faculty to think collectively about comprehensive end-of-program goals for student performance.

Portfolios have impressive face validity because they include tangible evidence of student performance over time. They reveal how students respond to a curriculum. And since students can select or create some of the contents of a portfolio, this assessment medium provides an excellent lens for observing individual differences in this regard. Portfolios not only stimulate faculty dialogue but also furnish a beginning point for faculty-student conversations. Moreover, when students use portfolios to present their achievements to potential employers, a third important kind of dialogue is initiated.

Portfolio use requires effective planning. The enterprise begins with goals for student learning, then students must be given specifications for portfolio content that will illustrate that the goals have been met. And faculty must agree on standards for judging the quality of the artifacts that are included. In Shay's experience, the exercise of setting these standards and then establishing that all who review the work have a common understanding of how to apply the standards (that is, achieving interrater agreement) can have the ancillary benefit of improving faculty attention to day-to-day assessment of student work in classes. Greater attention to the reliability of grading practices and to the need for giving students prompt feedback about their performance has been noted following training in portfolio assessment.

Assessment Update
March-April 1997
Volume 9, Number 2

Students must be carefully briefed about the purposes of their portfolios and the materials they are expected to include. As Spurr, Kiphart, and

Miller illustrate, effective portfolio evaluation depends on the presence of comparable materials across a group of students whose work is being evaluated.

Both Dillon and Spurr, Kiphart, and Miller speak to the several benefits of involving community representatives in portfolio evaluation. As society becomes more explicit in its expectations of higher education in terms of work force development, involvement of employers and other community representatives in setting standards for and then assessing student performance holds promise as a means of helping to meet these expectations. Those of us in urban settings have special appreciation for these possibilities.

In closing, I thank the colleagues whose work is presented in this issue for taking time to share their wisdom about portfolios with *Assessment Update* readers.

Students must be carefully briefed about the purposes of their portfolios and the materials they are expected to include.

Assessment Update
March-April 1997
Volume 9, Number 2

Bob Pace Tells Us What Students *Do* While in College

> *"Hi, I'm Bob Pace. The work you are doing at Tennessee reminds me a lot of what I was trying to do at Minnesota in the 1950s."*

As I recall, I first met Bob Pace in 1983. A colleague and I had just completed a presentation on our assessment program at the University of Tennessee, Knoxville (UTK), for a group of educators at a conference in Nashville. A gentleman of slight build, dressed in sport coat, pants, and shoes in pale shades appropriate for a warmer climate, made his way from the back row to speak to us. I will never forget his words, "Hi, I'm Bob Pace. The work you are doing at Tennessee reminds me a lot of what I was trying to do at Minnesota in the 1950s."

Pace went on to explain that during the years when he was engaged in institutional research, based at the University of Minnesota and later at Syracuse, that field was characterized by diversity and creativity. In studying the impact of college on student learning and development, Pace and his contemporaries in the 1930s, 1940s, and 1950s experimented with a wide variety of instruments, including diaries, time logs, personal essays; measures of interests, attitudes, and appreciation; alumni surveys; critical thinking and performance assessments; and various rating scales, in addition to the more traditional measures of general and special abilities and of generic and specialized academic achievement. Following this period of innovation and focus on student development, tests and measures became more standardized and the time of institutional research personnel was usurped by administrators asking for ever-increasing amounts of data for decision making and reports for state and federal agencies.

Pace saw in our work at UTK the possibility of a return to the variety and creativity of institutional research and the focus on student and teacher that had illuminated his career at Minnesota. I have such respect and admiration for Bob Pace that the impact of his compliment is not even eclipsed in my memory by the formal tribute accorded the assessment program at UTK the following year when it received the triennial award for outstanding use of measurement technology from the National Council on Measurement in Education.

Assessment Update
January-February 1996
Volume 8, Number 1

TRUDY W. BANTA

I do not often discuss specific assessment instruments in this column; Gary R. Pike is the expert in that arena. But since Pike's *Assessment Measures* column in this issue (page 296) features the Community College Student Experiences Questionnaire (CCSEQ), one of the family of Pace instruments developed during a remarkable career that spans seven decades, I want to say a little more about the technical quality, value, and utility of the CCSEQ and its companion for four-year institutions, the College Student Experiences Questionnaire (CSEQ).

Both the CCSEQ and the CSEQ are designed to assess the extent of students' involvement in desired learning activities in and outside the classroom. They are based on the well-documented premise that students learn more when they are actively engaged in their studies and in developmental activities that occur on campus. Both instruments also contain self-report measures of the progress students think they have made toward achieving specified education outcomes as a result of their college experiences. Student perceptions of various aspects of the college environment are probed in a third component of the two questionnaires.

Since he published the first version of the CSEQ in 1979, Pace has worked diligently to establish its reliability and validity (Pace, 1990). The same is true with regard to the CCSEQ. In addition, Pace has been perhaps our most influential scholar in providing evidence of the validity of student self-reports. And even though he retired from the University of California at Los Angeles, Pace is currently serving as institutional research adviser at Humboldt State University in northern California and continuing his research. In May 1995, he sent me a paper in which he argues convincingly that the CSEQ has yet another use: assessing the extent to which colleges encourage students to engage in contact with faculty, cooperation with other students, and active learning—all proven principles of good practice in promoting undergraduate achievement (Pace, 1995).

Students learn more when they are actively engaged in their studies and in developmental activities that occur on campus.

In gathering materials from well over 200 campuses for my own recent books on assessment, I can say that no other questionnaires with national norms have produced more positive change at those institutions than the CSEQ and the CCSEQ. For instance, at the University of Wisconsin-River Falls, where faculty allowed class time to be used by student affairs staff to administer the CSEQ to a third of all students, CSEQ data on student use of the library were used to support the case for a $7.1 million renovation of the library (Ballou, Reavill, and Schultz, 1994). A residence hall planning group used CSEQ findings to guide its work in refurbishing existing buildings and recommending new construction. A task force on the first-year experience found differences in CSEQ scores between first-year students who dropped out and those who persisted and thus was able to make responsive changes in programming.

At Santa Barbara City College, CCSEQ findings helped identify areas in which involvement was strongly related to student success and areas in which the institution needed to do more to encourage involvement (Friedlander and MacDougall, 1994). In-service training and mini-grants

Assessment Update
January-February 1996
Volume 8, Number 1

were provided to encourage faculty and staff to increase student involvement in classroom activities and participation in on-campus and community-based extracurricular activities. Counseling services were redesigned to nudge students into taking more responsibility for obtaining needed information. Faculty were urged to sponsor departmental clubs and out-of-class activities, and additional space for such gatherings was established in the Campus Center. A coffee house was created to increase student-student and student-faculty interactions.

Santa Barbara City College faculty and staff have been rewarded for their diligence in responding to CCSEQ findings: subsequent administrations of the questionnaire have shown gains in involvement, satisfaction, and progress toward desired college objectives.

The Tennessee Board of Regents sponsored a study in which the CCSEQ was administered to 633 students—176 African American and 457 Caucasian—enrolled in four west Tennessee community colleges to assess similarities and differences in the academic and social integration of these two groups of students (Friedlander, Murrell, and MacDougall, 1993). No noteworthy differences were found between the groups in quality of effort, academic engagement, or campus participation.

Santa Barbara City College faculty and staff have been rewarded for their diligence in responding to CCSEQ findings.

While Bob Pace continues his research on the instruments in California, the major responsibilities for marketing, scoring, interpreting, and improving the CSEQ and the CCSEQ have been transferred to academic centers in Indiana and Tennessee, respectively. George Kuh and his colleagues in the Center for Postsecondary Research and Planning at Indiana University now have the rights to the CSEQ, and Patricia Murrell and colleagues in the Center for the Study of Higher Education at the University of Memphis are in charge of the CCSEQ. Recently, a number of us convened to consider refinements in the CSEQ that will address the needs of the adult and commuter student populations at four-year urban campuses.

As Bob Pace (1995, p. 1) observed, "Good practices in undergraduate education . . . are usually described by what faculty and administrators do, but the existence of these good practices should also be evident in what students do." Pace's lifetime of good work in giving students avenues for telling us what they do while in college goes on-more solidly grounded now on campuses across the country than ever before.

References

Ballou, R. A., Reavill, L. K., and Schultz, B. L. "Administering the College Student Experiences Questionnaire: A Partnership Between Academic and Student Affairs." Unpublished manuscript, 1994.

Friedlander, J., and MacDougall, P. R., "Involvement." Unpublished manuscript, 1994.

Assessment Update
January-February 1996
Volume 8, Number 1

Friedlander, J., Murrell, P. H., and MacDougall, P. R. "The Community College Student Experiences Questionnaire." In T. W. Banta and Associates, *Making a Difference: Outcomes of a Decade of Assessment in Higher Education*. San Francisco: Jossey-Bass, 1993.

Pace, C. R. *The College Student Experiences Questionnaire*. Los Angeles: Center for the Study of Evaluation, University of California, 1990.

Pace, C. R. "From Good Processes to Good Products: Relating Good Practices in Undergraduate Education to Student Achievement." Paper presented at the annual meeting of the Association for Institutional Research Forum, Boston, May 1995.

Are We Making a Difference?

Has all of our activity—from state initiatives and accrediting requirements to assessment for improvement on campuses—made a difference?

Darrell Krueger, who is now president of Winona State University, was once dean of instruction at Northeast Missouri State University. Charles McClain, then president of Northeast, used to ask repeatedly, "Darrell, are we making a difference?" In seeking the answer, Krueger and McClain embarked on a comprehensive program of student outcomes assessment that transformed Northeast. From a teacher's college with projections of declining enrollment, it became a thriving regional university. Ultimately it was designated as Missouri's liberal arts and science institution.

Last year I undertook a national study designed to respond to McClain's question for the field of assessment. We are well into our second decade of outcomes assessment in higher education. Has all of our activity—from state initiatives and accrediting requirements to assessment for improvement on campuses—made a difference? Are students learning more? Are faculty teaching more effectively? Are colleges and universities better? In this column, I will provide some information related to this last question. I'll address the others in future issues of *Assessment Update*.

I have collected data from well over 100 campuses where outcomes assessment has been underway for several years-long enough to make some difference. In looking at the examples of impact, it seems to me that they can be grouped into at least three categories: *institutions* where assessment has transformed the campus culture; campuswide assessment *programs* that have resulted in significant, though less pervasive, improvements; and assessment *techniques* that have produced positive changes in specific areas. I will try to characterize the influence of assessment at some of the "transformed" institutions in the paragraphs that follow.

More than twenty years ago, the faculty at Alverno College employed an extensive program of assessment of their graduates to help reshape the curriculum around eight basic abilities: communication, analysis, problem solving, valuing in decisionmaking, interaction, global perspectives, effective citizenship, and aesthetic responsiveness. Since then Alverno faculty have

Assessment Update
September-October 1993
Volume 5, Number 5

developed ways to assess successive levels of these abilities in coursework, and their study of assessment has enabled them to improve course structure, teaching methods, and assessment techniques themselves. The quest of Alverno faculty for continuous improvement of student learning has enriched the understanding of faculty around the globe of such concepts as assessment as learning, contextual validity, portfolio assessment, assessment centers, and student self-assessment.

At Northeast Missouri, assessment was used initially to demonstrate that its graduates were "nationally competitive" in their chosen fields and that their experiences at Northeast had added value to their store of knowledge and skills. The university gradually acquired a national reputation for academic quality, which resulted in increases in both the number and the academic potential of students applying for admission. All these changes produced in the faculty at Northeast a sense of identity and pride in the institution that further transformed the university's culture.

Leaders at these transformed institutions have kept assessment at the forefront of everyone's thinking for years.

Strong presidents—Sister Joel Read at Alverno and Charles McClain at Northeast Missouri—focused faculty attention on assessment at their institutions in the 1970s. In 1985, another president—Nathan Weiss at Kean College—launched his own crusade to stimulate faculty self-consciousness about the quality of student learning using outcomes assessment. Kean faculty developed their own definition of assessment and with it a sense of power and responsibility for setting the direction for individual programs. They worked with each other, with students, with alumni, and with employers to define goals and objectives for their programs, to develop assessment strategies, and to make appropriate improvements. Kean College now has evidence that the student experience is richer and that student satisfaction has increased.

Presidential leadership was also an essential factor in development of the Institutional Impact Project at Ohio University in 1980. Testing, surveys, and student tracking have helped direct numerous improvements in programs and services at that institution. Assessment evidence now shows that students' scores on national exams have increased, freshmen are more involved in the campus experience, faculty and students interact more, seniors are more satisfied, and freshman retention has improved.

At the University of Wisconsin, Superior, strong administrative and faculty leadership and an assessment grant from the Fund for the Improvement of Postsecondary Education transformed a troubled institution with dangerously low enrollment to today's institution with rising enrollments and a much improved public image in its region. An assessment program that has helped to improve teaching, increase affective support for students, and raise levels of expectation and performance for students has produced more satisfied faculty, freshmen who report a more positive initial impression of the university, and students who find their classes more challenging.

Assessment Update
September-October 1993
Volume 5, Number 5

Leaders at these transformed institutions have kept assessment at the forefront of everyone's thinking for years. Assessment has been accorded time to develop in an orderly way, with a clear purpose, with widespread participation by all who should be concerned about the education of students: faculty, students themselves, administrators, graduates, employers, and community members. All five institutions have strong and effective assessment coordinators, drawn most often from the faculty. Since these individuals had earned the respect of their colleagues before they were given their assessment roles, they have been able to steer the campus assessment strategies through the sometimes treacherous shoals of faculty scrutiny and skepticism.

At these institutions, assessment findings are widely disseminated in a trusting atmosphere; negative results are not used to punish students or faculty. Assessment on these campuses is firmly anchored in established procedures that faculty care about, such as institutional planning and budgeting, curriculum development, faculty and staff development, and faculty scholarship. This fact, coupled with careful, persistent follow-up, has made assessment part of the fabric of these institutions and ensured its lasting influence.

Assessment findings are widely disseminated in a trusting atmosphere; negative results are not used to punish students or faculty.

The stories of assessment's impact on the five institutions just identified, as well as more than 100 others, have been collected in a book, *Making a Difference: Outcomes of a Decade of Assessment in Higher Education*, published by Jossey-Bass in October 1993. Because each chapter is written by the individuals who shepherded the assessment process being described, the volume has 45 coauthors. For the material I've just summarized, I'd like to thank the following authors: Georgine Loacker and Marcia Mentkowski at Alverno, Candace Young at Northeast Missouri, Michael Knight at Kean, Michael Williford and Gary Moden at Ohio University, and Albert Katz at University of Wisconsin, Superior.

Assessment Update
September-October 1993
Volume 5, Number 5

TRUDY W. BANTA

Does Assessment Help Faculty Teach More Effectively?

In my last column I posed three questions about the impact of assessment that 45 coauthors and I have attempted to answer in a new book, *Making a Difference: Outcomes of a Decade of Assessment in Higher Education* (San Francisco: Jossey-Bass, 1993). Those questions include: Are students learning more? Are faculty teaching more effectively? Are colleges and universities better? I intend to address these questions in reverse order in a series of three columns in *Assessment Update*. The question about colleges and universities was the subject of the Editor's Notes in the September-October 1993 issue. Here, I provide evidence that faculty are teaching more effectively.

Most faculty teaching in college today did not receive formal training in pedagogy as they prepared themselves in their disciplines.

Most faculty teaching in college today did not receive formal training in pedagogy as they prepared themselves in their disciplines. Thus, they have not studied theories of curriculum development, techniques for developing program goals and student learning objectives, strategies for helping students with a variety of learning styles get the most from a course, or methods for evaluating student learning. The apparent assumption underlying this state of affairs is either that this knowledge is not essential because a good scholar is, ipso facto, a good teacher or that teaching skills can simply be learned on the job.

Many conscientious faculty members who have observed that not all students profit from a steady diet of lectures—the most traditional and easiest to deliver mode of classroom presentation—have sought help in expanding and improving their repertoire of instructional techniques. But involvement in assessment has sent many more scurrying for guidance than would otherwise have done so. Moreover, they have approached professional development activities with more purpose and motivation because the acts of planning assessment activities and later interpreting assessment results have so clearly demonstrated the need for acquiring specific knowledge and skills.

For instance, faculty engaged in evaluating student portfolios at Miami University have come to see more clearly the interrelationships among

Assessment Update
November-December 1993
Volume 5, Number 6

BEST COPY AVAILABLE

courses in a curriculum, that is, that students are not just taking individual courses but an integrated curriculum. At Montclair State University, students' scores on a locally developed major field test revealed unexpected inconsistencies in the way in which faculty were teaching multiple sections of the same course. Both these sets of experiences have led faculty to focus attention on improving their skills in articulating curriculum goals and specific course objectives.

Faculty at many colleges and universities have been motivated via assessment to seek assistance in promoting student development of certain intellectual skills. At the State University of New York, Fredonia, locally developed tests of socioethical understanding produced results that led faculty to ask for and receive workshops on teaching problem-solving skills, scientific methods, and multicultural perspectives. Quantitative, humanities, and social science teaching methods have been strengthened at Empire State College as a result of faculty development stemming from the use of portfolios in student assessment. Low scores on the Communicating scale of the ACT-COMP examination caused faculty at Austin Peay State University to seek one another's counsel concerning ways to improve students' communication skills. A similar finding at the State Technical Institute at Memphis prompted the president to hire a specialist to assist all instructors in adding writing assignments to their courses. Formation and development of peer study groups has been the subject of professional development activity spurred by assessment at the University of Phoenix.

Faculty who take assessment seriously almost always come to question their proficiency in developing good measures of student learning in courses and programs.

Faculty who take assessment seriously almost always come to question their proficiency in developing good measures of student learning in courses and programs. Workshops aimed at helping faculty write better test items have been conducted at many institutions, including Johnson County Community College and the University of Tennessee, Knoxville. Colloquia on grading have been held at the University of Phoenix in response to the finding that the variance in course grades for adult students was excessively low, that is, that faculty were not making the full range of distinctions among levels of performance. Finally, the explosion of interest in using portfolios has created the need on many campuses for faculty to come together to learn how to score these collections of student work.

Because assessment has helped to increase faculty self-consciousness about teaching and concern about student learning, the chapter authors in *Making a Difference* have been able to provide a good deal of evidence that curriculum integration and teaching effectiveness have improved. Following professional development on the topic of liberal learning for adults, faculty at Empire State College have revised curriculum guidelines in order to make the meaning of general education clearer to students. A program evaluation process in the Eastern Iowa Community College District has resulted in specific curriculum or course changes in all vocational-technical programs to make the programs more responsive to identified needs of students and employers. Perceived weaknesses in the psychology major at Winthrop University identified in exit interviews with seniors have led faculty to add to the curriculum (1) a cornerstone course to make program objectives and

Assessment Update
November-December 1993
Volume 5, Number 6

options clear to incoming majors and (2) a capstone course designed to increase program integration.

Student tracking data at Seattle Central Community College revealed that only about half of the students enrolled in intermediate algebra courses achieved a grade of C or better. Subsequently, faculty revised this course to focus on graphing functions using real-world examples as topics for class discussion. The percentage of students receiving a C or better then increased from 53% to 71%. Interviews with entering students at the University of Virginia revealed such a high level of student dissatisfaction with the amount of faculty-student contact that a program of seminars for first- and second-year students taught by senior professors was undertaken.

The impact of assessment on student learning will be the subject of my next column. But I want to end this note with a hint of what is to come: at Miami University, faculty reading student portfolios have been able to document improvements in the use of scientific methods and quantitative reasoning following the increase in teaching emphasis that these skills have received in the course of responding to assessment findings at that institution.

I have just cited examples from thirteen institutions. A much fuller explanation of the impact of assessment on these campuses is the subject of 7 of the 24 chapters in *Making a Difference*. Knowing my colleagues as I do, I am certain that the 16 assessment coordinators who developed these 7 chapters will forgive me for not identifying them here; this recognition of the good work being accomplished at their institutions will satisfy them.

Are We Making a Difference in Student Learning?

It takes time to effect change in academic programs and methods of instruction and then chart the impact of these changes on students' knowledge and skills.

This is the third in a series of essays based on the content of the book *Making A Difference: Outcomes of a Decade of Assessment in Higher Education* (Banta and Associates, Jossey-Bass, 1993). In the September-October issue of *Assessment Update* (Vol. 5, No. 5, p. 3), I promised three articles, each addressing one of the following questions: as a result of a decade of intensive activity in outcomes assessment, (1) are colleges and universities better? (2) are faculty teaching more effectively? (3) are students learning more? The response to the first question appeared in Vol. 5, No. 5, and the second question was the subject of my column in Vol. 5, No. 6.

The short answer to the question about student learning is "probably so," but the quantity of concrete, objective evidence that assessment has improved learning is minuscule. I believe there are good reasons for this. First, it takes time to effect change in academic programs and methods of instruction and then chart the impact of these changes on students' knowledge and skills. With the exception of the half-dozen or so institutions that undertook comprehensive assessment programs before 1985, most institutions engaged in outcomes assessment have graduated no more than a class or two of students who might be expected to demonstrate higher levels of learning as a result of improvements made in response to assessment findings. Moreover, carefully planned longitudinal studies are needed to produce convincing evidence of change in students' knowledge and skills over time, and few institutions have invested the time and money necessary to conduct such studies.

Even where longitudinal studies have been undertaken, the inadequacies of available measurement instruments and methods have rendered the findings disappointing. For instance, at the University of Tennessee, Knoxville (UTK), where we conducted comparative studies for five years using the four best-known commercial tests of student achievement in basic knowledge and skills, we found that not one of those tests measured more than 30% of the content our faculty associated with general education. In addition, study after study revealed major technical flaws in the assessment of student score gain on these instruments (see Gary Pike's column in

Assessment Update
March-April 1994
Volume 6, Number 2

BEST COPY AVAILABLE

Assessment Update, Vol. 6, No. 2). Whereas most faculty would really like to observe actual behavior in order to assess performance, issues of reliability and validity, as well as cost, plague performance assessment as well.

Having pointed out the significant problems associated with providing the clear-cut quantitative data that would convince governors, legislators, trustees, parents, and students that assessment has improved student achievement, I am pleased to be able to cite some examples from a growing body of evidence that assessment is beginning to make a difference in student learning.

At institutions such as Alverno College, where faculty have agreed on a specific set of learning outcomes, each with its own progression of well-defined levels of achievement, and have then developed a curriculum based on those outcomes, improvements in student learning can be documented in each course and at the end of an academic program. In addition to experiencing faculty assessment of their development of eight abilities according to explicit criteria, Alverno students also continuously improve their capacities to assess their own performance. Alverno faculty analyzed the performance of alumnae and community members in work, personal, and citizenship roles as they defined their eight abilities and means of assessing them. A continuing longitudinal study of alumnae both confirms the value of the learning experience at Alverno and suggests directions for improvement of curriculum, instruction, and methods of assessment.

I am pleased to be able to cite some examples from a growing body of evidence that assessment is beginning to make a difference in student learning.

At Northeast Missouri State University, turning student and faculty attention to the need to be "nationally competitive," as measured by scores on nationally standardized tests, initially had a dramatic impact. Scores on the ACT-COMP exam increased, especially after a new math requirement was instituted in response to faculty disappointment with student scores on the quantitative portions of the test. Longitudinal studies also showed score improvement over time on the COMP as well as on standardized tests in various majors. Continued charting of student progress on the same tests has proved problematic, however, because the nature of the tests and their norming populations have changed over the years and because Northeast's success has attracted a brighter and brighter student body, whose performance on standardized tests is both qualitatively and quantitatively different from that of Northeast's student population in the 1970s.

Ohio University has reported steady increases in student scores on the ACT-COMP exam since 1980. COMP scores of seniors in nursing at Austin Peay State University increased after faculty modified their curriculum and advising procedures to improve students' experiences in communicating and valuing.

At the University of Wisconsin-Superior (UWS), seniors' scores on the Major Field Achievement Test in Psychology increased after faculty instituted more demanding course requirements in areas of relatively poor student test performance. Dyersburg State Community College (DSCC) is one of many institutions where specific pedagogical interventions have produced improvements in the percentage of graduates passing the licensing exam in nursing.

Assessment Update
March-April 1994
Volume 6, Number 2

Faculty at the State Technical Institute at Memphis (STIM) are among those at a majority of institutions in believing that their students' writing skills need improvement. At STIM a writing specialist was hired to help all instructors incorporate writing assignments in their classes and to ensure that tutoring services in communication were available to all students. Objective assessment of students' lab reports and other writing assignments indicates that these responses to assessment findings have produced improvement in writing skills. Similar efforts to improve writing have had the desired impact at Northeast Missouri, Kean College, UTK, and Winthrop University. At Miami University, increased faculty emphasis on scientific methods and quantitative reasoning has produced improvements in the written work submitted in portfolios in these areas.

Concern about the high percentage of students failing algebra at Seattle Central Community College led faculty to begin teaching graphing using graphing calculators and to create student groups for the purpose of studying equations. These changes in instructional practices apparently have increased the percentage of students receiving a grade of C or higher from 53% to 71%.

> *Faculty can point to a number of indirect indicators which at least suggest that students are learning more where assessment is conscientiously implemented.*

While we work to improve the technical quality of direct measures of student achievement, faculty can point to a number of indirect indicators which at least suggest that students are learning more where assessment is conscientiously implemented. At Northeast Missouri, students are spending more time studying and more time using library materials. Longitudinal research at Ohio University indicates that students are spending more time with faculty, both formally (as in advising) and informally (as in having dinner together). Ohio graduates responding to a survey say that their college experience has increased significantly their abilities to communicate orally and in writing, evaluate and choose between alternative courses of action, think analytically, and apply knowledge learned in their major. At UWS, 40% of students surveyed recently said that the university's assessment program had made their classes more challenging.

While persistence and graduation statistics do not in themselves indicate enhanced levels of achievement, the very fact that a student stays in college longer and takes more courses strongly suggests that he or she has learned more. At Ohio University, freshman-to-sophomore retention increased from 67% to 85% between 1978 and 1990. DSCC pushed its fall-to-spring retention rate from 60% to 67% and increased by 60% between 1989 and 1992.

Finally, improvements in the placement of graduates that have occurred at a number of institutions as a result of specific changes made in response to assessment findings may be considered indicative of improved student learning. At Ohio University more than 90% of the graduates are now employed or in graduate school at the end of the year following their graduation. And at Northeast Missouri, the percentage of graduates going on to graduate or professional education tripled between 1975 and 1990—from 10% to 30%.

Assessment Update
March-April 1994
Volume 6, Number 2

TRUDY W. BANTA

Now We Are Ten

Welcome to the tenth season of *Assessment Update*! In 1988 Gale Erlandson, higher education editor at Jossey-Bass, asked if I would be willing to edit that publisher's first newsletter. With substantial support from my colleagues at the University of Tennessee, Knoxville, I agreed to take on the project. Nine years later, despite the pressure of meeting a deadline every other month and coping with the necessity of reading each article we publish at least four times, I continue to enjoy the responsibility! It has been gratifying to watch the manuscripts increase in number and diversity of content over the years. And since my move to Indiana in 1992, I have enjoyed the full support and competent assistance of colleagues at Indiana University-Purdue University Indianapolis.

> ***This informal publication provides a history of the development of the field and related issues at the classroom, departmental, campus, state, regional, national, and international levels.***

In preparing for our tenth year, we have worked with Jossey-Bass staff to develop a new look for *Assessment Update*. We hope you will find it more attractive and inviting to open and easier to read. (If you have comments—pro or con—about the new format, please send them to me.) In addition, we have recently added a Web Corner feature that will keep you abreast of assessment developments that are as accessible as your computer screen.

Also in anticipation of our tenth anniversary, we have developed a collection of all the columns contributed over the years by each of our four featured columnists—Peter Ewell, Peter Gray, Gary Pike, and Jeff Seybert—plus my own *Editor's Notes*. A number of people have requested such a collection over the years, and putting it together has given each of us an opportunity to reflect on developments in assessment over the past decade. This informal publication provides a history of the development of the field and related issues at the classroom, departmental, campus, state, regional, national, and international levels.

In this issue, Peter Ewell and Peter Gray provide their first columns for Volume 10. Ewell describes some of the effects on higher education of the deregulatory policy environment that developed in Washington following the 1994 congressional elections, then focuses on advances in student tracking

Assessment Update
January-February 1998
Volume 10, Number 1

BEST COPY AVAILABLE

that technology is making possible. Peter Gray's *Campus Profile* is about Ball State University in Muncie, Indiana. I have long admired the work of Catherine Palomba, who coordinates Ball State's assessment program. The multiple-measures approach at Ball State incorporates a judicious combination of standardized and locally developed measures of student learning and satisfaction.

As a student in one of my graduate seminars, Sheila Ewing wrote an extensive review of the literature on performance assessment. She included so many citations and organized them so effectively that I urged her to adapt the paper for *Assessment Update*.

Peter Kugel presents a wonderfully straightforward and logical approach to assessment that has, fortunately, been successful at several research universities—that is, identifying some questions of interest to a group of faculty, gathering some data, and acting on the findings. Anne Hummer reports on a faculty-developed instrument for assessing students' perceptions of a fundamental attribute: their communication skills.

We hope to make the tenth volume of *Assessment Update* the best ever. Please send your own suggestions for ensuring that we achieve our goal.

Assessment Update
January-February 1998
Volume 10, Number 1

TRUDY W. BANTA

Benchmarking in Assessment

Benchmarking. Now there's a word that has come into its own in the 1990s! As happens with any term that describes a popular concept, people have offered many definitions. I think of benchmarking as a process of standard setting—looking for exemplary practice in a well-defined area, investigating carefully the components of such practice, distilling some generalizations, and trying to apply that learning to improve one's own practice.

Until recently, we in higher education have done our own benchmarking more or less informally. If our institution is a member of an organization like the American Association of Universities or the Urban 13, we may work with colleagues in those organizations to develop some common indicators that help us judge our competence in attracting the best-prepared students or retaining students of color—just two of hundreds of measures in which we might be interested. To bring the study even closer to home, many of us have selected ten or so institutions we consider peers and have begun to exchange information with them about such matters as faculty salaries and levels of student satisfaction with instruction and campus services. A few disciplinary associations such as those in business and law have also served as conduits for exchange of comparative data.

Until recently, we in higher education have done our own benchmarking more or less informally.

But sharing data is only one small step in the benchmarking process. It is certainly an important step, because colleagues can't share information for long without beginning to ask each other *exactly* how data are gathered and how statistics are defined. Such conversations usually produce efforts to standardize data definitions and some of the processes that produce the data. However, standardizing measures can take place at a distance from the practices that produce the measures. Only recently have we begun to get together to investigate carefully the components of good practice.

In 1996, Peter Ewell introduced me to the work of the American Productivity & Quality Center (APQC), a nonprofit organization based in Houston. Peter had been invited to serve as a subject matter consultant to APQC for a study designed to identify and disseminate exemplary practice in measuring

Assessment Update
July-August 1998
Volume 10, Number 4

BEST COPY AVAILABLE

institutional performance outcomes. Ultimately my institution, Indiana University-Purdue University Indianapolis, joined the consortium of institutions engaged in that study. Because we subsequently were selected as one of the "best practice institutions" for the in-depth investigation component of benchmarking, I have had an opportunity to learn firsthand how APQC's benchmarking studies take place. Moreover, I have been asked to serve informally as one of the subject matter consultants on APQC's current study on assessing learning outcomes.

In 1992, APQC and eighty-six of the country's quality-oriented companies designed the International Benchmarking Clearinghouse to help managers find and adopt exemplary practices with the aim of improving their performance. Since 1993 this clearinghouse has conducted over thirty consortium benchmarking studies. APQC's subsidiary Institute for Education Best Practices has thus far carried out benchmarking studies of institutional budgeting, electronic student services, and faculty development, as well as the two studies on measuring institutional performance outcomes and assessing learning outcomes.

An APQC benchmarking study begins with the identification of an area for study and one or more subject matter experts like Peter Ewell. Then APQC staff, with assistance from the consultant(s), develop a list of potential sponsors for the study. Selected two- and four-year higher education institutions, corporations, and health care or nonprofit agencies were contacted about participation in the Assessing Learning Outcomes project. More than thirty colleges and universities, corporations, and federal agencies ultimately paid a fee and signed on as sponsors.

I have had an opportunity to learn firsthand how APQC's benchmarking studies take place.

Following a pre-planning phase, a kickoff meeting for representatives of sponsoring organizations was held at APQC headquarters in Houston in November 1997. At that meeting the sponsors refined the scope of the benchmarking study, nominated potential "best practice partners," created criteria for screening these partners, and discussed potential questions for a screening survey.

Following the November meeting, APQC staff, assisted by the content expert, developed a screening survey draft that was subsequently sent to sponsors for critique. They then sent a revised survey to potential best practice partners.

APQC staff compiled the survey data they received and mailed a summary for each responding institution or organization to sponsors. In February 1998, a review meeting of sponsors took place in Houston. Through a series of consensus-building techniques, the group identified six best practice partners. The list included Ball State and Emporia State Universities, Sinclair Community College, the University of Phoenix, Fidelity Investments, and TVA University.

At the February meeting, sponsors also developed a draft of the questions they would like to have answered during visits to the best practice partners.

Assessment Update
July-August 1998
Volume 10, Number 4

Throughout the spring, sponsors had the opportunity to take part in preplanned visits to the six sites of exemplary practice.

Subsequently, the APQC staff compiled summaries of the best practice partners' responses to the standard set of questions posed by the sponsors during the site visits. In June both partners and sponsors attended a sharing session in Houston. Each best practice partner made a brief presentation and responded to additional questions from sponsors. The content specialist and APQC staff identified key findings of the benchmarking study, including innovative practices and critical success factors and enablers.

Following the June meeting, the next—most difficult—steps in the benchmarking process will be up to the representatives of the sponsoring organizations. To capitalize on their investment, they must take the most appropriate parts of what they learned in Houston and from the site visits, adapt these findings to their own settings, and implement the actions that are designed to improve their own practice.

APQC's approach to benchmarking is new, interesting, and potentially of value to those of us who believe our processes in higher education should be improved continuously. Readers of *Assessment Update* who want to learn more about APQC and its studies may consult the Web site <www.iebp.apqc.org> or contact APQC's Marisa M. Brown at (800) 776-9676 or <mbrown@apqc.org>.

APQC's approach to benchmarking is new, interesting, and potentially of value to those of us who believe our processes in higher education should be improved continuously.

That Summer Reading List

How do we engage faculty in assessment?

I am so glad that Barbara Walvoord and her colleagues chose *Assessment Update* as the place to publish "Closing the Feedback Loop in Classroom-Based Assessment." They are leading us toward a very powerful response to that all important question, How do we engage faculty in assessment?

In the article, Walvoord, Barbara Bardes, and Janice Denton describe how faculty can agree on some broad learning outcomes, then collectively submit the following for departmental and campus-wide discussion: (1) classroom tests or assignments that show how they have examined students on those outcomes, (2) detailed scoring rubrics in the form of Primary Trait Analyses (see *Assessment Update* 7:6 and 9:3), (3) aggregated student scores using the rubrics, and (4) a description of ways in which faculty will use assessment findings derived from the rubrics to improve student learning of the specified outcomes. Thus, in assessment faculty use tests and assignments they have developed and rubrics that reflect their own criteria and standards. Because they derive students' grades from the assessment, both faculty and students are motivated to take it seriously. And taking the final crucial step-closing the feedback loop by looking at students' performance in the aggregate and determining where and what action may be needed to address lower than hoped for scores-seems relatively easy and sensible.

In *Effective Grading: A Tool for Learning and Assessment* (Jossey-Bass, 1998), Walvoord and Virginia Johnson Anderson have given us additional information about how to improve assessment in that most important of all sites, the classroom, where improvement can have the most direct and immediate impact. They suggest that once faculty identify what they most want their students to know and do in a course, the next step should be to look carefully at assignments and tests to see whether they both teach and assess the knowledge and skills identified in the first step. Next, a course outline can be constructed that puts the assignments and tests in a logical sequence. Then the instructor can decide what needs to happen in each class to prepare the students to gain the most from the assignments and tests. Finally, the authors explain how Primary Trait Analysis can be used

Assessment Update
September-October 1998
Volume 10, Number 5

BEST COPY AVAILABLE

to give more precise meaning to the scoring of assignments and tests, and ultimately to course grades.

Summer will be long gone by the time you read this. But describing Walvoord's exciting work reminds me of some other materials I have come across during the summer that I hope you will come to value as I do.

Elizabeth Jones, who has left the National Center on Postsecondary Teaching, Learning, and Assessment at Penn State and is now a faculty member in educational policy and leadership at West Virginia University, has written in these pages (*Assessment Update* 5:6, 8:6, and 9:6) about the studies on skills identification that she has conducted during the past four years with support from the U.S. Department of Education's Office of Educational Research and Improvement. As a first step toward assessing progress toward the national goal that addresses college student learning, Jones worked with broadly representative panels of college faculty, employers, and state and national policymakers to identify and define the essential skills named in National Goal 6, Objective 5: communicating, critical thinking, and problem solving. Faculty interested in assessing these skills in their classes and curricula will find helpful the detailed descriptions of student outcomes contained in a series of goals inventories that Jones has developed in writing, critical reading, speech communication, critical thinking, and problem solving. To learn more, contact Beth Jones at <ejones7@wvu.edu>.

Describing Walvoord's exciting work reminds me of some other materials I have come across during the summer that I hope you will come to value as I do.

An additional resource for classroom or disciplinary assessment is the work on reflective thinking and problem solving being done by Cindy L. Lynch, independent scholar, and colleagues at a variety of institutions (see, for example, the "Campus Strategies" column by Phillip Wood and Cindy Lynch in *Assessment Update* 10:2, pp. 14-15). Courses in which faculty are assessing complex thinking and professional problem solving with Cindy's assistance include critical reading at California State Fullerton, electrical engineering at the U.S. Air Force Academy, and accounting at the University of Denver. For definitions of terms used in these assessments (such as identifying, framing, resolving, and re-addressing), a detailed evaluation rubric, and examples of exercises designed by faculty to teach and assess problem solving, see *A Developmental Guide to Assessing and Optimizing Professional Problem Solving* (Lynch, Wolcott, and Huber, 1998) at Susan Colcott's Website <http:/www.du.edu/~swolcott>. For additional information, contact Cindy Lynch at <Leehaven@compuserve.com>.

Faculty interested in portfolio assessment will enjoy the brochure entitled *Your Portfolio: Liberal Arts and Sciences* which was developed for students by faculty at Truman State University. Under the subheading, "Why Should I Maintain a Portfolio?" the student will find detailed comments from recent graduates describing what the process of developing a portfolio meant to them-both before and after graduation. Other subheadings include "How Do I Start and Maintain a Portfolio?", "How Can I Use the Portfolio?", and "How Might the University Use My Portfolio?" For more on Truman's portfolio assessment process, contact Ian Lindevald or Shirley

Assessment Update
September-October 1998
Volume 10, Number 5

Morahan, codirectors of portfolio assessment, at <lindy@truman.edu> or <smorahan@truman.edu>.

Steven J. Osterlind is a careful, competent scholar at the University of Missouri-Columbia who has devoted more than a decade to the development of the College Basic Academic Subjects Examination (College BASE), a criterion-referenced achievement test that assesses skills and competences usually associated with the completion of general education coursework. Now Steve has analyzed the sources of some 73,000 students at fifty-six colleges and universities who took the College BASE between 1988 and 1995 and has written *A National Review of Scholastic Achievement in General Education* (ASHE-ERIC Higher Education Report, Vol. 25, No. 8, 1997). Among Osterlind's disturbing findings is the fact that 20 percent of the very large number of students tested—most of whom had completed their coursework in general education—proved to be poor readers or weak writers, or both. Only 25 percent were able readers and just 3 percent were skilled writers. Faculty, administrators, and policymakers seeking benchmarks against which to assess generic skills and knowledge of college students will find Osterlind's report essential reading.

Only by seeing our enterprise as a system and by working on many fronts simultaneously can we hope to have an impact on the dismal portrait of collegians' achievement.

Finally, let me emphasize the importance of rereading, as I did this summer, Peter Ewell's article, "Organizing for Learning: A New Imperative" in the December 1997 *AAHE Bulletin*. Ewell warns that change initiatives like general education reform, outcomes assessment, service learning, and technology-enhanced instruction undertaken piecemeal, in isolation from one another and from the most valued processes for accomplishing change within our institutions, will not enjoy long-term success. He summarizes some essential lessons from the fields of cognitive science, educational psychology and instructional design, and organizational change and suggests how these research findings can be integrated into a systematic approach to improving teaching and learning in the academy. Only by seeing our enterprise as a system and by working on many fronts simultaneously can we hope to have an impact on the dismal portrait of collegians' achievement that Osterlind has painted for us.

Assessment Update
September-October 1998
Volume 10, Number 5

Section III. Assessment in Response to External Pressures: Performance Funding in Tennessee

Having been drawn into outcomes assessment in 1979 as the University of Tennessee, Knoxville (UTK) attempted to respond to the Tennessee Higher Education Commission's performance funding initiative, I could not resist writing about this unique program in my column from time to time. My first description of the program appeared in 1990. As much as we chafed under the requirements to use standardized instruments and be graded on the basis of students' responses, by that time I had witnessed sufficient new attention to the learning environment for students on the part of faculty engaged in assessment to be willing to argue that external pressures for accountability COULD influence faculty in positive ways (see 2:1).

External pressures for accountability COULD influence faculty in positive ways.

In 4:2 I presented the other face of performance funding: the negative influences associated with Tennessee's use of estimated scores on standardized tests to calculate—and reward—the value added by the college experience.

By 1994 I was moved to write positively about performance funding again (6:4). I like to think that the research we conducted at UTK from 1986 to 1992 had some impact on the improvement of the performance funding guidelines promulgated at the state level in 1991. A study we conducted in 1992-93 enabled us to write a thorough critique of the performance funding criteria, and this study is summarized in 6:4.

BEST COPY AVAILABLE

TRUDY W. BANTA

A Small Controversy

For a full year—through four issues—I have resisted the temptation to devote the space for a lead article to a description of our work at the University of Tennessee, Knoxville (UTK). *Assessment Update* is a national publication and will not be used to promote any particular program, model, method, or point of view at the expense of any other. Nevertheless, I am proud of the accomplishments of my faculty and administrative colleagues at UTK, and the opportunity to provide a counterpoint to Tom Jeavons's thesis with the UTK experience as an illustration, was simply irresistible.

The financial incentive to conduct outcomes assessment activities is so substantial that no institution can afford to ignore it, and none does.

No other state-level outcomes assessment initiative is as broadly prescriptive as that developed by the Tennessee Higher Education Commission (THEC). Publicly supported institutions are directed to seek accreditation for every accreditable program (or show cause why this should not be done); to test every senior in general education, using a specific standardized exam; to test seniors periodically in every major, using specific standardized exams when these are available; to administer a specific alumni survey biennially; and to use the results of these assessment activities in taking corrective action designed to improve academic programs and student services. Complex scoring criteria are applied in connection with each of the activities, and an institution's total score determines the proportion it receives of a budget supplement equivalent to 5.45 percent of its state allocation for instruction.

The stakes are high indeed for Tennessee institutions that are subject to the THEC performance funding program. The financial incentive to conduct outcomes assessment activities is so substantial that no institution can afford to ignore it, and none does.

If faculty in Tennessee, by adapting assessment methods to serve their own purposes and by using findings to improve the learning environment for students, can respond to an external mandate as intrusive as THEC's performance funding, then it is possible for any faculty to do so. This is not to say

Assessment Update
Spring 1990
Volume 2, Number 1

BEST COPY AVAILABLE

that every faculty member will respond positively; many will not, but it is difficult to name any issue on which all faculty agree.

I do not believe an external mandate is the ideal stimulus for involvement in assessment. Fortunate is the faculty that perceives a need and chooses outcomes assessment as a means of addressing it. In the absence of this happy circumstance, however, we can still derive benefit from the outside pressures we face.

Thomas H. Jeavons's Reply

Whatever controversy may exist between Trudy Banta and myself is cast in shades of gray, not black and white. We both recognize that assessment gains faculty support only when it can be shaped by faculty members to serve what they see as significant educational purposes. Trudy and I differ in our perceptions of the approaches likely to serve such purposes and in our sense of the probability that these approaches will be adopted when the agenda for assessment is initially established outside the academy.

Fortunate is the faculty that perceives a need and chooses outcomes assessment as a means of addressing it.

It is not clear to me how so-called performance funding is implemented in Tennessee. But in many states across the country, faculty members are worried—with good reason—that the performance measures being sought will be used by those outside academe to punish the programs that are doing less well, not to provide extra resources to those who need them. Moreover, when such measures fail to consider differences in the kinds of students coming into particular programs or differences in the missions of institutions, this sort of assessment for comparative purposes is truly invidious. Where faculty members perceive a push to "quantify" performance in a way that ignores variations in circumstances, no assessment program is likely to win faculty members' acceptance or engage their creativity and become the kind of positive experience that Trudy describes at UTK.

Frankly, many assessment mandates do not appear to allow the kind of flexibility and creativity that has occurred at UTK. I am impressed by much of what Trudy describes at UTK, but I remain deeply skeptical about the ability of standardized tests to assess the capacities that we hope a liberal education will generate. The students participating in the New England economics project in our study were the sort that do very well on the GRE (a commonly cited endpoint assessment for undergraduates), but they did not show the intellectual agility that faculty expected them to demonstrate near the end of their education. These faculty members wanted to know whether their students were attaining crucial abilities to integrate different kinds of learning, think creatively, work collaboratively, and apply general competencies to specific problems—the most significant fruits of a solid liberal education. Such capacities will not be discerned by paper-and-pencil tests; they will be discerned only by faculty engaged personally with their students in situations where those capacities are put to work.

Assessment Update
Spring 1990
Volume 2, Number 1

This leads to my last point: that locally designed, creative approaches to assessment can at the same time yield useful information about students' learning and lead students to new learning. Such approaches can generate activities that will immediately enhance students' education, rather than simply evaluate it. And if we are concerned about making the best use of scarce resources, these are the kinds of options we need most to be seeking.

We welcome your responses to our respective arguments.

Creative approaches to assessment can at the same time yield useful information about students' learning and lead students to new learning.

Assessment Update
Spring 1990
Volume 2, Number 1

Some Unvarnished Truth

Successful programs begin with goals and specific objectives, are implemented with strategies tailored to achieving the objectives, and are evaluated with measures closely connected to the implementing processes.

In 1984, in a volume he called *The Self-Regarding Institution* (published by NCHEMS), Peter Ewell identified three examples of the self-regarding genre: Alverno College, Northeast Missouri State University, and the University of Tennessee, Knoxville. Despite the fact that faculty at Northeast and UTK have developed most of the measures used in their comprehensive approaches to assessment, both institutions have been perceived within the assessment community as relying heavily on the results of standardized tests. In the mid-1970s, Northeast faculty and administrators made decisions about the kinds of evidence needed to answer President McClain's question, "Are we making a difference?" In the early 1980s, UTK faculty had to decide how to respond to a directive from the Tennessee Higher Education Commission that seniors be tested to qualify the institution for performance funding. At the time, standardized exams seemed to both groups to provide an alternative worth exploring.

In Virginia, New Jersey, and most other states, the mandate to assess student outcomes has been less prescriptive than in Tennessee. And as will be clear to regular readers of *Assessment Update*, most faculties—almost 70%, according to the 1991 ACE Campus Trends Survey—have elected to develop many of their own measurement instruments.

In the lead article for this issue, Darrell W. Krueger, formerly the chief architect of Northeast's award-winning assessment program and now using a different blueprint for assessment at Winona, gives us some good reasons for the interest in local initiatives. Successful programs begin with goals and specific objectives, are implemented with strategies tailored to achieving the objectives, and are evaluated with measures closely connected to the implementing processes. While standardized tests can focus faculty attention on goals and may yield some of the evaluative information we seek, the diversity in American higher education virtually ensures that no single standardized instrument will match exactly the specific objectives of any given faculty-designed program. Moreover, as Krueger points out, the test scores themselves do not even suggest what processes need to be improved to further

Assessment Update
March-April 1992
Volume 4, Number 2

BEST COPY AVAILABLE

the achievement of program goals. For Winona, the Wingspread Principles are providing direction for process improvement, and TQM is keeping everyone's attention on the goal of using assessment results to effect continuous improvement.

Darrell Krueger's article begins with his disillusionment with test scores as bases for improvement initiatives at Northeast but ends with an exploration of more hopeful alternatives. I hope this piece provides an antidote for the somewhat acrimonious debate between Joe M. Steele of ACT and my own UTK colleague, Gary R. Pike, which occupies much of the rest of this issue.

In 1982, when a UTK faculty committee recommended that we try the ACT-COMP exam to assess student learning in general education, COMP was the only standardized test available for this purpose. Our alternatives were either to give the Graduate Record Exam or to readminister the ACT Assessment EXAM given to entering students, both of which were designed for very different purposes. Committee members were not persuaded that the COMP provided a good match with faculty goals for general education—essentially a set of distribution requirements—and they were very skeptical about the ACT staff proposal to estimate freshman-to-senior gain scores. The committee directed its campus assessment coordinator to do three things that were not required by the state but were in keeping with our responsibility as Tennessee's research university: (1) test freshmen as well as seniors, so that we could see for ourselves how ACT's estimates compared with our own reality, (2) conduct studies on the psychometric properties of estimated and actual score gain, and (3) establish a program for systematically comparing the COMP exam with other instruments that might be developed in the future for assessing general education programs.

Accountability as the raison d'etre of assessment is becoming stronger than ever in today's economic climate.

With the support of UTK administration, we began to carry out this set of directives in 1983. We moved very carefully, cautiously, deliberately, sharing our findings with ACT staff and asking for their reactions before we made any public statements. As Gary Pike indicates, our first article on estimated gain was not published until 1987; the first comparative studies appeared in 1989; and now we have a report on actual freshman-to-senior gains.

Some will fault me for devoting so much space to a technical matter that is a bread-and-butter issue in only one state—Tennessee alone bases state funds on estimated score gain. But the "value-added" issue is far from dead. As Peter Ewell keeps telling us in his *From the States* column and elsewhere, accountability as the raison d'etre of assessment is becoming stronger than ever in today's economic climate, and more legislators and public interest groups are asking, "What are students learning in college?" For many, the quickest, simplest answer is to be found in a gain score derived from a standardized test.

Steele and Pike lay out the issues in the debate over score gain. We invite you to join the discussion. Write to either of the authors, preferably both, and/or to me. I would like to publish thoughtful commentary on the subject in this space in the future.

Assessment Update
March-April 1992
Volume 4, Number 2

Performance Funding in Tennessee Comes of Age

How effective are these externally imposed accountability measures in stimulating internal improvement initiatives?

Assessment Update
July-August 1994
Volume 6, Number 4

Since its inception in 1979, performance funding in Tennessee has been one of the most widely observed and discussed of all the state approaches to improving institutional quality in higher education. (See Peter Ewell's column, "Performance Funding: New Variations on a Theme" in *Assessment Update* Vol. 6, No. 4.) Over the fifteen years since the first financial incentives for evidence of student learning were proposed by the Tennessee Higher Education Commission (THEC), three-quarters of the other states have issued calls for institutions to demonstrate their accountability for the use of public funds; but no other state or agency has based its accountability program on specific test scores and numerical satisfaction ratings to the extent that THEC has done.

In 1994 Tennessee institutions are completing the second year of the state's third five-year plan for performance funding. This new iteration of THEC's policy—the fourth revision since the first guidelines were developed in 1979—includes a much more finely tuned, campus-responsive, and comprehensive set of criteria than the set that has guided the process previously. While test scores and satisfaction ratings are still prominent criteria of quality in the new guidelines, methods for determining the statistical significance of differences between annual campus means and established norms have now been developed and are being applied in determining institutional scores. Moreover, substantial weight is now accorded peer judgment in the form of program reviews for general education, unaccredited undergraduate, and master's-level programs. And, finally, institutions now may seek part of their awards on the basis of meeting their own targets for increasing the diversity of their student populations and for making progress toward other mission-specific goals.

Are these new performance funding guidelines considered more effective by faculty and administrators than those in effect during the 1980s? How is each of the current funding criteria perceived on the campuses as a measure of the quality of higher education? How effective are these externally imposed accountability measures in stimulating internal improvement initiatives?

BEST COPY AVAILABLE

Three former Tennessee colleagues and I undertook a study designed to provide answers to these questions. Linda Rudolph of Austin Peay State University, Homer Fisher from the University of Tennessee, Janice Van Dyke at the State Technical Institute in Memphis, and I developed a detailed questionnaire asking for a critique of each of the standards and sent it to the assessment coordinator at each of Tennessee's 23 public colleges and universities. By virtue of diligent follow-up, we achieved a response rate of 100%.

We first asked the assessment coordinators to evaluate each of the ten performance funding standards "as a measure of the quality of higher education" by assigning a letter grade—A, B, C, D, or F—using the same system they would apply in judging students' work. Despite the fact that more than one-half of the ten criteria have been used as measures of quality in Tennessee for fifteen years, no standard was accorded an average grade of A by these experienced assessors.

Five standards were given grades of B. The highest marks went to three standards that involve peer review: peer review of unaccreditable undergraduate programs, peer review of master's-level programs, and accreditation. (For two-year institutions, placement of graduates is substituted for peer review of master's degree programs.) A fourth standard, improvement actions designed to utilize assessment results in improving programs and services, also received a grade of B, as did student and alumni satisfaction surveys. The five standards given C ratings included testing in major fields, progress toward mission-specific goals, testing in general education, progress in meeting retention and graduation goals, and progress on enrollment goals for minorities and other special groups.

We also asked the assessment coordinators to respond to the question "Has this standard led to improvement in student learning on your campus?" Here, accreditation of accreditable programs headed the list as a stimulus for improvement: 85% of the respondents believed that this standard, which has been in effect since 1979, has improved student learning on campus.

At least two-thirds of the survey respondents reported that master's-level reviews and placement of graduates, improvement actions, and surveys—also performance funding standards with long histories—have actually produced improvements in student learning. More than one-half of the respondents were willing to say this about major field testing (55%) and peer review of undergraduate programs (52%). The latter rating is all the more noteworthy because it belongs to the only new (in 1992) standard that made it into the top tier for improvement potential.

Underlying the last question on the survey is the one asked most often around the globe with regard to assessment: Can assessment undertaken to meet accountability demands produce internal improvement on campuses? Our questionnaire asked respondents to think about all of the standards together and to rate the potential of the entire performance funding program for improving institutional effectiveness. The overall rating given by the assessment coordinators was C+.

Our questionnaire asked respondents to think about all of the standards together and to rate the potential of the entire performance funding program for improving institutional effectiveness.

Assessment Update
July-August 1994
Volume 6, Number 4

BEST COPY AVAILABLE

Collectively, Tennessee's assessment coordinators have more years of experience than any other group of academics in the world in implementing campus assessment initiatives in response to externally imposed mandates. They and their faculty colleagues were consulted extensively during the nearly two-year process used by the THEC staff to revise the standards for the third five-year plan. Why, one might ask, does there seem to be so little enthusiasm for performance funding in Tennessee after fifteen years of state and campus involvement in trying to get it right?

We are drawn back to something leaders in our field have been saying for years: assessment simply goes against the grain for academics. First, we are not sure that we can or should specify what students ought to learn as a result of their college experiences. In fact, we know that the most valuable thing they will learn is to make their own meaning of their experience. Second, we are not satisfied with the measures of student learning that are available to us; we recognize that their reliability and validity are far from perfect. We resent the time that must be spent on assessment; it takes time away from teaching and talking informally outside class with students, activities that we know have a positive effect on student learning. Moreover, our involvement in assessment is not sufficiently recognized or rewarded when promotion and tenure decisions are made.

Assessment simply goes against the grain for academics.

As Thomas Angelo says in the lead article of *Assessment Update* (Vol. 6, No. 4), "If assessment is ever to improve substantively the quality of student learning . . . both faculty and students must become actively, continuously, and personally involved." In Tennessee, as in every other state, too few faculty and students have become personally involved in assessment, even when large sums of money are based on it.

My coinvestigators from Tennessee believe that the overall grade of C for the performance funding standards is overly harsh. They point out that our study took place in the second year of a five-year program—too early to judge the ultimate impact of the addition of three new standards that only reward institutions for progress on long-term goals. They say that the written comments about the effectiveness of the individual standards reveal more positive attitudes than the assigned letter grades convey.

On the positive side, we can certainly say that after fifteen years of experience with assessment for accountability purposes in Tennessee, a majority of the academics who have observed performance funding activities most closely are willing to conclude that the accountability measures of peer review and surveys for constituent groups such as students, alumni, and employers, as well as a planned program of improvement actions based on assessment findings, are assessment methods that also improve student learning on campus. Moreover, the view from the campuses is that the act of giving institutions the opportunity to receive additional state funds for setting their own goals and methods of monitoring in areas that are of interest to the state, such as increasing the minority enrollment and retention and graduation rates, also holds promise for serving the twin purposes of accountability and improvement.

Assessment Update
July-August 1994
Volume 6, Number 4

Section IV.
Some National Assessment Issues

In 1989 I wrote in 1:3 that Peter Ewell had said at the Fourth National Conference on Assessment in Higher Education, "Assessment is not a profession, nor should it become one." I added, "Thus there are no experts, just people sharing ideas and learning from one another." I went on to say that "Assessment practitioners can expect to find guidance in the literature of educational measurement, program evaluation, policy analysis, institutional research, strategic planning, organizational development, research utilization, and organizational change, to name some of the most obvious areas."

In the next issue (1:4) Marcia Mentkowski joined me as a co-author of *Editor's Notes*. The column was entitled "Collaborating in Setting Directions for Assessment Research" and ended with a challenge to "interested researchers:" "We stand ready and willing to identify issues and problems gleaned from our experience. Data are available for study. Most important, we are eager to review your findings and use them to build more effective assessment strategies."

One of my greatest frustrations is that ten years later there still are very few experts and virtually no "interested researchers" in any of the fields in my long list from 1989 who have analyzed our data and helped to ground our experience in theory that would lead to hypothesis-testing. Thus the assessment community remains largely a community of "just people sharing ideas and learning from one another." As I said in the column entitled "The Rising Star of Assessment Scholarship" (8:6, which is included in the *Methods* section of this review) there is an encouraging increase in action research, or reflection on practice, among assessment practitioners in a variety of disciplines. But assessment is far from the point of being strengthened by a discernible connection with, for example, theories of change, organizational development, or research utilization. Fortunately, this may be just on the verge of changing. In 1998, in issue 10:2, which appears in this volume in the section I've headed "Encouraging Faculty Involvement . . ." I summarize some of the points in a December 1997 *AAHE Bulletin* article by Peter

There is an encouraging increase in action research, or reflection on practice, among assessment practitioners in a variety of disciplines.

BEST COPY AVAILABLE

Ewell that suggest how assessment can and should be linked with ongoing research on metacognitive development, educational psychology, instructional design, continuous improvement, and organizational development.

Columns 2:3 and 3:5 describe my experience as a member of the Special Study Panel on Education Indicators that was appointed in 1989 by the Secretary of Education to advise the National Center on Education Statistics (NCES) on the development of an appropriate set of indicators for assessing the health of the education system in the U.S. The Study Panel released its report, *Education Counts*, in September 1991 (3:5).

In 3:6 I described the research program of the National Center on Postsecondary Teaching, Learning, and Assessment, which was established at Penn State in 1990.

It was clear that the states were not willing to carry the measurement baton any farther down the road.

In 1992 I began to write about my experiences with the NCES initiative to further National Goal 5.5 (later 6.5) related to assessing college students' abilities to think critically, communicate effectively, and solve problems (4:1). I prepared one of 15 papers that initiated the discussion about whether and how this assessment should take place. In 1993 I reported on another phase of this debate (5:4) and in 1994 I outlined the proposal of Dennis Jones and Peter Ewell to substitute, for the time being, indirect indicators of good practice for the direct measures of college student learning that seemed to most of us to be too costly and difficult to develop and administer (6:1). Finally, in 8:2 I reported on the last NCES-supported conference on this National Goal. By the time of the conference in December 1995 hopes of NCES staff for funding the development of a national assessment of college student learning had evaporated with the withdrawal of Congressional support. The purpose of the December meeting was to see if NCES might help states identify ways to measure the identified abilities. By the end of the conference it was clear that the states were not willing to carry the measurement baton any farther down the road. I advanced a modest proposal for encouraging regional accreditors to develop indicators for judging the effectiveness of colleges and universities in promoting student learning, then incorporating annual reporting on these indicators as a component of the widely respected process of peer review.

TRUDY W. BANTA

A Community of Learners

Ted Marchese's "Adding It Up" session at the conclusion of the fourth National Conference on Assessment in Higher Education (Atlanta, June 1989) provided the opportunity to reflect on the conference. Some of my own impressions may mirror those of others in the audience.

Peter Ewell mused, "It pleases me that this conference is a community of amateurs. Assessment is not a profession, nor should it become one." Thus there are no experts, just people sharing ideas and learning from one another.

And how eager we were to learn from one another! Pat Hutchings and I wondered if anyone would show up for our open forum, "Motivating Students for Assessment," at 8:30 a.m. on Saturday, the fourth day of the conference. What began as a crowd of well over 100 at 8:30 quickly grew to standing room only, and seldom had we seen such lively table-talk following our structuring of the discussion. Interesting ideas emerged, including the following:

> ***We wondered if anyone would show up for our open forum, "Motivating Students for Assessment," at 8:30 a.m. . . . What began as a crowd of well over 100 at 8:30 quickly grew to standing room only.***

Communicate expectations about assessment to parents as well as students, especially for freshmen.

Ask students how assessment activities can benefit them personally, and then use what they say in publicity designed to inform other students.

In addition to telling individuals about their performance, give the student newspaper and radio station periodic reports on selected campus-wide assessment results.

More faculty than ever before attended this fourth national conference. The first conference, in 1985, attracted many chief academic officers, deans, and representatives of accrediting associations. The sessions focused primarily on responses to the question "Why should institutions undertake assessment?" Now that almost 70% of the nation's colleges and universities have

Assessment Update
Fall 1989
Volume 1, Number 3

BEST COPY AVAILABLE

some assessment activity underway (see *Memos* for more from the American Council on Education's *Campus Trends*, 1989, survey), far more faculty are involved, and the predominant question in Atlanta was "How can we do assessment effectively?"

About a dozen of us who have been close to campus assessment programs for most of this decade had some opportunities in Atlanta for conversation about the future of assessment. With assistance from the Fund for the Improvement of Postsecondary Education (FIPSE), some of us have gathered on several previous occasions for extended consideration of selected issues. The lead article, by Stephen B. Dunbar, on the implications of K-12 competency testing for assessment in higher education, is an outgrowth of one such session.

In Atlanta we talked about the need for developing theory to guide our work and for drawing on established lines of inquiry for this purpose. Assessment practitioners can expect to find guidance in the literature of educational measurement, program evaluation, policy analysis, institutional research, strategic planning, organizational development, research utilization, and organizational change, to name some of the most obvious areas.

While some hope to blaze new trails in assessment, we must never forget the needs of beginners in our field.

While some hope to blaze new trails in assessment, we must never forget the needs of beginners in our field. One of the conferees took the audience microphone during the closing session to say, "I've been here for four days, and I still don't know exactly how to get started on my own campus." A panelist replied, "Find some other faculty members who share your interest in improving instruction, and begin by asking yourselves, 'What do we want students to learn, and how will we know if they've learned it?' The rest will follow."

Assessment Update
Fall 1989
Volume 1, Number 3

TRUDY W. BANTA

Collaborating in Setting Directions for Assessment Research

As assessment gains momentum, a growing number of practitioners is working together to influence the direction of the enterprise. Consortia, conferences, computer networks, and this newsletter—all are aimed at providing collaborative leadership, direction, and dialogue.

Currently, assessment strategies are improved on the basis of day-to-day feedback of assessment information to individual students, departments, or institutions. Practitioners are gradually stepping back from immediate involvement to consider questions that may improve practice even more, if pursued in more depth and over a longer term. Forming these questions collaboratively as an assessment community is an important step toward attracting the interest of colleagues from the research community in making an investment in assessment and beginning a dialogue about what to study.

Forming research questions collaboratively is an important step toward attracting colleagues' interest in assessment.

An ongoing activity that demonstrates such collaboration and leadership is the American Association for Higher Education (AAHE) Research Forum. Organized in 1985 as a way for educators to identify research questions in higher education, the Research Forum is a regular feature of each AAHE annual meeting. Conference presenters and participants engage in a specially designed group process over two days, to generate a set of questions that represent new issues emerging over the past year's work. Educators intend these questions to serve as a starting point for research aimed at improving practice. (For details of the process, see M. Mentkowski and A. Chickering's article in the Winter 1987 issue of *The Review of Higher Education*, "Linking Educators and Researchers in Setting a Research Agenda for Higher Education.")

From the start, session leaders Arthur Chickering (George Mason University), K. Patricia Cross (University of California, Berkeley), Catherine Marienau (DePaul University), and Marcia Mentkowski (Alverno College) have recognized the need to devote Research Forum topic areas to assessment. This year, interest in assessment commanded two of the Forum's six discussion topics: "Assessment of Student Learning"

Assessment Update
Winter 1989
Volume 1, Number 4

BEST COPY AVAILABLE

and "Assessment of Institutional Effectiveness." Guided by group leaders Jean McGregor (Evergreen State College), Trudy Banta (University of Tennessee, Knoxville), and Donald Lumsden and Henry Ross (Kean College of New Jersey), 13 conference presenters generated an initial set of questions. Then 37 participants at a second session added their critiques and observations, to shape a final research agenda on assessment. As part of a continuing annual tradition, this agenda was then disseminated to members of the American Educational Research Association.

More than 50 research questions were developed. Examples related to assessment of student learning include the following: how does each of the successful assessment models promote student development? How can assessment activities of faculty, student development personnel, and others be linked effectively? What kinds of performance feedback to individual students stimulate them to improve? Does this differ for the beginning student and the advanced student? In what ways can students build the capacity for self-assessment throughout their lives? How can we communicate the complexities and limitations of assessment to audiences outside the academy?

We are eager to review your findings and use them to build more effective assessment strategies.

Questions concerning assessment of institutional effectiveness include the following: What will be the impact of assessment on student diversity? What will be the impact of standardized tests on minorities? What kind of assessment program is viable for transfer students? What discipline-specific assessment strategies can help us determine the impact of teaching on learning? How can the effectiveness of menu-style general education programs be assessed, given the variability of students' course selections? How can we track the effects of diverse courses on writing, speaking, and critical thinking across the curriculum? What constitutes institutional readiness for developing a successful assessment program? How can varied modes of reporting be integrated to communicate assessment results to different constituencies?

The full set of questions from the 1989 AAHE research agenda, "Improving the Odds for Student Achievement: A Research Agenda," as well as prior agendas and reprints, are available from Alverno College's Office of Research and Evaluation, 3401 S. 39th St., Milwaukee, WI 63215.

The collaborative efforts of this widely representative body of assessment practitioners send this message to interested researchers: we stand ready and willing to identify issues and problems gleaned from our experience. Data are available for study. Most important, we are eager to review your findings and use them to build more effective assessment strategies.

Assessment Update
Winter 1989
Volume 1, Number 4

TRUDY W. BANTA

Our Readers Speak (We Wish You'd Write!)

We were pleased that Jossey-Bass included an evaluation form with the third issue of *Assessment Update* (*AU*). Responses were received from only 10% of our subscribers, but those who did complete the form provided a wealth of good suggestions.

Eighty-three percent of these readers said *Assessment Update* meets their professional needs "moderately" or "very well," 16% responded "occasionally," and 1% said "not at all." Ninety-two percent find the overall appearance of the newsletter attractive, 78% said the layout makes it easy for them to find the information they need, 71% believe *Assessment Update* provides the latest information in the field, 67% find the feature articles substantive and important, 62% consider *Assessment Update* practical and useful, and 58% agree that the newsletter's coverage "adequately reflects the scope and diversity of assessment efforts around the country."

Seventy-one percent believe* Assessment Update *provides the latest information in the field.

The feature articles garnered the most positive response: 90% of the respondents rated them good (45%) or excellent (45%). The regularly appearing signed columns contributed by Peter Ewell, Peter Gray, Gary Pike, and Trudy Banta received a response of good or excellent from 86% to 89% of the respondents. No section of the newsletter had more than 14% of the responses in the fair-to-poor category.

The quarterly publication of *Assessment Update* is acceptable to 68% of the respondents; 31% would like to receive the newsletter every other month, and 1% would prefer it twice annually. Almost 60% think the field is ready to support a full-length journal. Unfortunately, 54% said that few of their colleagues in the field know about *Assessment Update*; 38% said some know, and only 8% believe that most of their colleagues know about the newsletter. A promotional mailing from Jossey-Bass was the chief source of initial information about the publication.

When asked what they liked most about *Assessment Update*, readers identified the newsletter format and the fact that it is timely, brief, and easy to read and

Assessment Update
Summer 1990
Volume 2, Number 2

BEST COPY AVAILABLE

comprehend. They commended the breadth and variety of coverage and appreciate having information about assessment on campuses all over the country included in a single publication.

However, the brevity that is considered a virtue by some is a frustration for others. The chief complaint about *Assessment Update* is that it is too brief and that the articles are superficial in their coverage. (Nevertheless, respondents applauding brevity outnumber its detractors by three to one.)

Survey respondents offered many helpful suggestions. The most frequent request was for more information about assessment at community colleges. We have added Kay McClenney, who is a popular consultant in the two-year sector, as a consulting editor. She contributed a feature article for our fourth issue. She will help to keep us abreast of events, issues, and innovations in community college assessment.

The biggest single need of readers seems to be for practical, "how-to-do-it" information about assessment strategies.

The biggest single need of readers seems to be for practical, "how-to-do-it" information about assessment strategies—how to select or design instruments, how to solicit faculty support and motivate students to participate, how to ensure the use of results. We hope that we are addressing that need in all sections of the newsletter. We pledge to be mindful of the plea "Don't forget those involved in assessment by appointment rather than by training or experience; help educate us to make good administrative decisions regarding assessment."

This initial response from our readers reinforces our belief that faculty and administrators in higher education are ready to read and write about assessment. We are glad to find so many readers in the field, but frankly, we wish more of you were accepting the responsibility of becoming *writers* for *Assessment Update*. At the moment, we publish almost 90% of the material we receive, and we invite most of the authors of the articles we return to submit the same material in a much abbreviated form. We are pleased with the quality and usefulness of the articles you are sending, but Margery Bensey and I would feel more secure if we were not operating so close to the margin.

A majority of the articles sent to *Assessment Update* comes from just five states: Virginia, Missouri, Tennessee, New York, and New Jersey. As more states acquire track records in assessment, we hope to hear from campuses in many other states.

A personal letter may be the easiest way for you to communicate. Please take the time to write me a brief letter about a meeting you've attended, a helpful resource you've read, or an assessment problem you've solved. You may well find that information shared with colleagues in a forthcoming issue of *Assessment Update*.

Assessment Update
Summer 1990
Volume 2, Number 2

TRUDY W. BANTA

Education Indicators

In September, 1989, Secretary of Education Lauro Cavazos fulfilled a provision of the Hawkins-Stafford Elementary and Secondary School Improvement Amendments of 1988 by appointing the 19-member Special Study Panel on Education Indicators. The amendments direct the National Center for Education Statistics (NCES) "to issue regular public reports to the President and Congress on dropout and retention rates, results of education, supply and demand of teachers and school personnel, libraries, financial aid and on other education indicators." The Special Study Panel is to make recommendations concerning the kinds of education indicators that NCES should collect and report.

I confess that I regarded my appointment to the panel with decidedly mixed feelings.

Secretary Cavazos has construed the legislative mandate very broadly and directed the Special Study Panel to consider the need for preschool and postsecondary indicators, as well as those associated with elementary and secondary education. The panel includes economists and business leaders, teachers and professors, educational consultants, and state as well as local school administrators.

On June 28, 1990, in Washington, D.C., the panel, under the leadership of its chair, Alan Morgan, chief state school officer of New Mexico, began its third two-day meeting. I would like to take this opportunity to inform the higher education assessment community of the progress of the panel's work.

I confess that I regarded my appointment to the panel with decidedly mixed feelings. I was honored to be invited to join a federal advisory committee, especially one that might produce a report with the impact of *Involvement in Learning* or *A Nation at Risk*. Nevertheless, the topic was "indicators," a dry statistical term and one that I hoped would not become associated with a vast new set of reporting requirements for higher education. I took my oath as a panel member, vowing privately to hold the line against such a development.

Assessment Update
Fall 1990
Volume 2, Number 3

BEST COPY AVAILABLE

In six days of meetings with the entire panel in Washington, and in two days with special interest groups in Chicago, some of my initial fears were dispelled. I have developed enormous respect for my fellow panelists; each is a true expert in his or her own field. Most also share my concern that resources spent on data collection may not produce educational improvements.

The panel has adopted a very broad view of its task. Members have recognized the legitimate fear that a singular focus on assessing outcomes may cause us to ignore the processes that produce them, and may obscure clues to solving the problems that assessment suggests.

Panelists have agreed on six areas in which indicators may be classified. Three are related to the context of education, and two describe what goes on in schools themselves; only one is exclusively outcomes-oriented.

The indicators must be significant, enduring, and easily interpretable.

The context areas include society's support for schools and learning (business and community support and learning opportunities), readiness for school (children's health and access to preschool), and educational equity for students at risk for educational and social failure (access, by race and gender, to educational opportunities).

The school-related areas concern (1) the quality of schools and educational experiences (administrators' leadership and instructors' preparation) and (2) appreciation of, engagement in, and acquisition of subject matter, advanced academic and thinking skills, and competence for citizenship.

The outcomes-oriented area deals with educational contributions to economic productivity (quality of the workforce and international competitiveness).

In meetings that will take place over the coming year, the Special Study Panel will identify, for each of the six areas, indicators that NCES can use to guide its data collection activities. The indicators must be significant, enduring, and easily interpretable.

Assessment Update
Fall 1990
Volume 2, Number 3

TRUDY W. BANTA

Education Counts

Selection of national indicators can define the national agenda for education. What we decide to measure will dictate what we work to improve.

My column in the Fall 1990 issue of *Assessment Update* was entitled "Education Indicators." My purpose then was to provide an interim report on the work of the Special Study Panel on Education Indicators appointed in 1989 by the Secretary of Education. I'd like to summarize some of the highlights of the panel's final report, *Education Counts: An Indicator System to Monitor the Nation's Educational Health*, released in 1991.

The panel was charged with the task of examining models of educational systems, criteria for selecting educational indicators, and related data needs and collection efforts, then recommending for the National Center for Education Statistics (NCES) an indicator development plan and new research and data collection activities. Panel members represented business leaders, teachers, principals, parents, chief state school officers, school superintendents, and academics from an array of disciplines.

Within the past year (1991), several developments have served to heighten the importance of the panel's work. Former President Bush and the nation's governors agreed on six goals for education and began to talk about means of assessing progress, and the President and then Education Secretary Lamar Alexander announced a national education strategy, "AMERICA 2000," for moving the country toward those goals. A national consensus has developed about the need to improve schools and colleges and to raise the achievement levels of graduates. The report, *Education Counts*, suggests how the Federal government can establish the means for assessing progress toward the current set of national goals. Moreover, it recommends that the United States develop a comprehensive indicator information system that is sufficient to enable us to monitor the health of the education enterprise well into the next century.

Members of the indicators panel recognize that the selection of national indicators can define the national agenda for education. What we decide to measure will dictate what we work to improve. Thus the panel rejected the notion of developing a small number of key education indicators. Instead,

Assessment Update
September-October 1991
Volume 3, Number 5

BEST COPY AVAILABLE

we defined six major issue areas that we considered significant and of enduring educational importance.

The panel considered and rejected the input-process-output and general goals-specific objectives-measurement models of indicator development in favor of the broader issues approach. Members hope that this system will incorporate enough fundamental ideas, priorities, and concepts to encourage the public to see the interconnectedness of the education enterprise. By encompassing both schools and colleges in its recommendations, the panel underscores the importance of the connections among all sectors of education.

While the panel believes that a comprehensive indicator system should include measures of student learning, its report asserts that a system built solely on achievement testing is insufficient and will mislead the American public. An indicator system must go beyond assessments of student performance to include descriptions of the educational environment—teacher quality, curriculum, school organization, community support, and equity. If students are not learning, we need to know *why*, and test scores alone will not give us the answers.

We defined six major issue areas that we considered significant and of enduring educational importance.

The panel has established three criteria for a comprehensive indicator system, each of which is addressed by two of the six issue areas identified. The issue areas are consistent with the goals identified by the President and the governors in 1990, but they go well beyond those goals to broaden the vision of what the public should be concerned about as it seeks to improve its schools and colleges.

The first criterion is that indicator information must focus on what matters most about learning and about schools and colleges. The two issue areas that help to fulfill this criterion are *learner outcomes*; acquisition of knowledge, skills, and dispositions; and *quality of educational institutions*.

The second criterion is that indicator information must assess trends that bear on education and learning, today and in the future. *Readiness for school* and *societal support for learning* are the issue areas most closely related to this standard.

Finally, indicator information must reflect the nation's values and collective aspirations for education. The relevant issue areas here are *education and economic productivity* and *equity*: resources, demographics, and students at risk.

Education Counts contains suggestions for clusters of indicators that might be developed in each of the six issue areas. For example, the area *quality of educational institutions* subsumes four primary concepts: learning opportunities, teachers and teaching, schools as places of purpose and character, and school resources. Within the *learning opportunities* category, indicators could be designed to describe exposure to subject matter, nature of learning opportunities, assignment of teachers and students, and curricular integration.

Assessment Update
September-October 1991
Volume 3, Number 5

The panel envisions an "information pyramid" for each issue area in which composite or key indicators are composed of clusters of indicators that are, in turn, composed of selected statistics drawn from research findings and profiles gathered through data collection. Its report provides specific guidance only at the composite and cluster levels. The panel recommends that NCES be restructured to provide periodic reports on progress in each of the six issue areas as well as interpretive reports that integrate information across issue areas on selected topics of pressing importance, such as national economic competitiveness. The panel also recommends that NCES strengthen its system of advisory groups, using representatives of schools and colleges, the business community, and the public to assist in identifying statistics and indicators, directing data collection efforts, and developing interpretive reports.

Assessment Update
September-October 1991
Volume 3, Number 5

The National Center on Postsecondary Teaching, Learning, and Assessment

The principal purpose of the national center is to discover what facilitates student learning.

No field can long endure in the absence of a sound basis in fundamental research. Outcomes assessment in higher education was launched in the 1980s largely as an applied field. The external pressure to produce information for accountability purposes did not permit institutions the time to engage in serious reflection, much less the kind of experimentation that could give assessment programs a solid grounding in established lines of inquiry. While the new field has found sufficient support to grow and even flourish at some institutions, it cannot be sustained without a carefully constructed basic and applied research program.

In 1990, the National Center on Postsecondary Teaching, Learning, and Assessment was established to provide leadership in developing the research base for our field. The center, based at Pennsylvania State University, is supported by a five-year grant from the U.S. Department of Education's Office of Educational Research and Improvement. Its work is planned and carried out by faculty and administrators at six institutions, including the University of Illinois at Chicago, Syracuse University, Northwestern University, Arizona State University, and the University of Tennessee. The center is co-directed by James L. Ratcliff and Patrick T. Terenzini.

The principal purpose of the national center is to discover what facilitates student learning. It also plans to conduct studies designed to improve the educational effectiveness of institutional, state, and federal policies and practices. These inquiries will be conducted through a longitudinal panel study and four research programs. The panel study and research programs are described here, with particular emphasis on their relationship to the field of outcomes assessment.

A sample of approximately 5,000 undergraduates from more than 20 colleges and universities will be followed for three years, beginning with the freshman year. The purposes of this panel study are (1) to estimate how academic and nonacademic experiences influence student learning, attitudes toward learning, cognitive development, and persistence in

Assessment Update
November-December 1991
Volume 3, Number 6

BEST COPY AVAILABLE

college and (2) to determine the extent to which these experiences differ among the various kinds of students who attend the variety of colleges and universities in this country. The *academic* experiences to be studied include the quality of teaching and classroom instruction, the level of student involvement in academic work, and the pattern of coursework taken. *Nonacademic* experiences include relationships with faculty and peers, extracurricular activities, and work and family responsibilities.

Research Program on the Curriculum. Using current knowledge about goals, purposes, and standards of the undergraduate curriculum, several studies have been designed (1) to explore the content and cognitive abilities and the values and norms needed by students to achieve these goals, and (2) to increase understanding of how teachers can transform that knowledge base into effective curricula. Two programs in this area, "Indicators of Learning Within the Major" and "Effective Patterns of Coursework in General Learning," examine the basic and critical thinking skills required in the undergraduate curriculum. A third study, "Effect of Coursework on Student Learning," considers how the common curricular experience of students affects their learning and development.

Research Program on Faculty and Instruction. Three research projects in this program have direct connections with our work in outcomes assessment. The "Enhancing Teaching" project includes several studies on the gathering and use of information by faculty to assess and improve instruction. The "Instructional Methods and Minority Students" project is using research on learning and teaching styles to identify instructional methods compatible with the ways minority students learn in community colleges. The "Collaborative Learning" project examines the effects of collaborative learning techniques on teaching and students.

Research Program on Out-of-Class Experiences. Component projects in this area focus on how students become active and involved participants in an academic community and how their out-of-class experiences, particularly their on- and off-campus interpersonal interactions, reinforce or weaken curricular and classroom learning and achievement of broader general education goals.

Research Program on Organizational Structures and Policies. This research program is designed to identify administrative approaches to improving undergraduate teaching, learning, and assessment. Areas of inquiry include (1) organizational characteristics that advance or hinder innovation in teaching and learning, (2) information and data essential to effective decisionmaking about teaching and learning, and (3) leadership activities necessary to create a climate for change.

For information, contact Maryellen Weimer, Associate Director, 403 South Allen St., Suite 104, University Park, PA 16801-5202. Tel: (814) 865-5917; fax: (814) 865-3638.

Assessment Update
November-December 1991
Volume 3, Number 6

Take Part in the National Goals Debate

Experienced assessment practitioners know that we must also be concerned about achieving consensus on the goals of assessment and on the learning outcomes to be measured.

Former President Bush, Secretary of Education Lamar Alexander, and the nation's governors agreed on six national goals for education that provide the basis for the *America 2000* program of educational improvement. While this program is aimed principally at the K-12 sector, the involvement of higher education is essential in preparing teachers, upgrading workforce skills, and shedding light on the best methods of promoting and assessing student development. In addition, National Goal 5, Objective 5, touches directly the interests of those of us concerned about assessment in higher education. That objective is stated, "[By the year 2000] the proportion of college graduates who demonstrate an advanced ability to think critically, communicate effectively, and solve problems will increase substantially."

The first annual progress report on the achievement of the national goals, which was issued on September 30, 1991, did not include any information on achievement of Objective 5.5. However, its section "Indicators for Future Reports" contained an expression of hope that indicators could be developed for the abilities of college students to think critically, communicate effectively, and solve problems.

Members of the National Education Goals Panel and others involved in guiding implementation of *America 2000* strategies are advocating the use of multiple measures in any national assessment system that may be developed; that is, multiple-choice tests would be supplemented with performance-based or portfolio assessments. But the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education have issued a joint statement urging caution in moving to base a national system on these relatively new and unproven technologies without considerable exploration of the complex technical issues involved.

As Donald Stewart points out in the lead article in this issue, experienced assessment practitioners know that we must also be concerned about achieving consensus on the goals of assessment and on the learning

Assessment Update
January-February 1992
Volume 4, Number 1

outcomes to be measured. What effects will the very process of trying to reach this consensus in higher education have on the scope of college curricula, on academic freedom, on institutional diversity?

We also know that students must be motivated to do their best work if assessment is to be valid. What consequences will be associated with student performance on a national examination system? The response to this question may have the effect of limiting the opportunity to obtain a college degree and ultimately restricting access to higher education.

Finally, we are painfully aware that assessment in and of itself does not produce educational improvements. Assessment must be linked to specific outcome-oriented objectives, which are systematically implemented with appropriate human and technical resources. Only if these connections are explicit can assessment findings be used effectively to suggest directions for improvement.

A decade of widespread experience in assessing the outcomes of higher education has produced a number of important principles, which should be considered in the debate surrounding design and implementation of a national assessment system, especially one that includes a postsecondary component. We have learned that faculty, students, and other stakeholders must be involved in setting the goals and objectives for learning and assessment; that these objectives must be used in the classroom to shape teaching and student development; that students must accept responsibility for their own learning and performance on assessment measures; that we must employ a variety of assessment techniques; and perhaps most importantly, that we must build assessment into ongoing processes that matter to faculty, so that they will use assessment findings to improve curricula and methods of instruction. Now is the time to share what we have learned with decisionmakers who hope to improve education in this country.

Now is the time to share what we have learned with decisionmakers who hope to improve education in this country.

Find out what the governor of your state is doing to promote the *America 2000* goals, and volunteer to serve on task forces and advisory groups—even if the current focus of their activity is K-12 education. You can contribute to early discussions about reforms at that level, thus establishing credibility as a resource when the debate turns to assessment in higher education.

Sal Corallo reported in the November-December 1991 issue of *Assessment Update* that the National Center for Education Statistics (NCES) has asked 15 individuals with a diverse array of scholarly interests, from the definition of critical thinking to the evaluation of new response modes for standardized tests, to contribute position papers on the implementation of Objective 5.5. The operative word here is *implementation*—the writers were instructed to raise questions and concerns if they wished but to give NCES the benefit of their best thinking about how the objective should be operationalized, because it *will* be implemented in one way or another.

Assessment Update
January-February 1992
Volume 4, Number 1

The NCES panel of 15 includes 5 people with particular interest in higher education assessment: Georgine Loacker and Marcia Mentkowski of Alverno College, Peter Ewell of the National Center for Higher Education Management Systems, James Ratcliff from Penn State, and myself. Any and all of us would be pleased to hear how you are becoming involved in the conversation about the national goals. I would be delighted to publish your letters on this topic or to quote from them in this space.

Assessment Update
January-February 1992
Volume 4, Number 1

TRUDY W. BANTA

Do Faculty Sense the Tightening of the Accountability Noose?

Academic year 1992-93 was sobering for those of us who hope that conscientious implementation of outcomes assessment can be a significant factor in enabling colleges and universities to demonstrate their worth to external constituencies. Ever increasing competition between higher education and a diverse array of social services for ever scarcer public funds and among a diverse array of providers for commitment from those shopping for educational services has led to more intense external scrutiny of our institutions than ever before.

To provide a sampling of the forms this scrutiny has taken: the public perception that research activities are diverting faculty time and attention from undergraduate teaching has led legislatures and governing boards to undertake inquiries into how faculty are spending their time. Regional and disciplinary associations have become more diligent in seeking evidence that institutions are seriously engaged in assessing outcomes. A task force appointed by the National Governors' Association (NGA) recommended implementing a national performance-based assessment for a sample of postsecondary students to measure progress toward the national goal (National Goal 5, Objective 5, to be precise) of increasing "the proportion of college graduates who demonstrate an advanced ability to think critically, communicate effectively, and solve problems." And the National Center for Education Statistics (NCES) has issued a request for proposals to initiate the first phase of this national assessment.

I have found myself right in the middle of all of this. Within three months of my arrival, my new institution was reviewed by the North Central Association and was subsequently instructed to develop by January 1995 a formal plan for assessing student learning. The trustees of Indiana University asked us for a report on the nature of faculty work and the means being used to evaluate it. The NGA invited me to join Peter Ewell and measurement expert Stephen Dunbar from the University of Iowa in a public debate with the chair of its task force on postsecondary assessment. Despite my strong reservations about the development of a national test, I agreed to contribute ideas to a consortium of organizations submitting a proposal to NCES for the first phase of this development.

> ***I have found myself right in the middle of all of this.***

Assessment Update
July-August 1993
Volume 5, Number 4

BEST COPY AVAILABLE

We have ample evidence, especially from the experience of elementary and secondary schools over the last quarter century, that testing alone does not improve learning. In fact, far from having produced exemplary levels of learning, achievement testing in grades K-12 now reveals some serious deficits in fundamental skills for U.S. students as compared with their age peers in many other countries. These findings are among the most influential reasons for the growing interest in restructuring local schools and school systems.

The development of performance-based assessment for a sample of postsecondary students will be enormously expensive and will not in and of itself improve the learning of college students. Realization of Goal 5, Objective 5 requires that (1) the skills of communicating, problem solving, and critical thinking at a level acceptable for college graduates can be operationally defined, (2) a substantial proportion of U.S. faculty will agree with the definitions and teach the skills in ways that are effective in promoting student learning of them, (3) valid measures of the defined skills can be devised, (4) a truly representative sample of students can be induced to participate in the national assessment and to take this exercise seriously and perform conscientiously, and (5) the assessment results will be used to improve teaching and other aspects of the educational experience for undergraduates.

We have ample evidence . . . that testing alone does not improve learning.

Once again drawing on the experience of our K-12 colleagues, we know that this sequence of activities can best be carried out at the local level in individual colleges and universities and in individual classrooms. Why do we have to spend years of precious time and millions of dollars of public funds to find this out?

Perhaps a large expenditure of time and money is necessary if only to call faculty and administrators' attention to the public's concern about the outcomes of higher education.

The NGA has sponsored a series of hearings on the issue of national postsecondary assessment. Despite the fact that they have been scheduled in conjunction with the annual meetings of several large national organizations that have a substantial number of members from academia, the hearings have not drawn large audiences. For example, the testimony heard by NGA in Atlanta in one of the hotels hosting the American Educational Research Association (AERA) conference, which annually attracts six to eight thousand people, drew no more than 20 AERA members for the NGA audience.

The accountability noose is being drawn ever more tightly, and ultimately this fact will become clear to faculty throughout academe. As more voices are heard in opposition to the national postsecondary assessment plan and the actual costs of the enterprise become known to those who must fund it, this particular train may well be derailed. But in the meantime, NCES is making sure that the train leaves the station on the first leg of its journey—the objective-setting phase—and I hope some of us who have the benefit of campus-based experience in outcomes assessment are at least on board ready to give advice to the engineer about the curves and obstacles ahead.

Assessment Update
July-August 1993
Volume 5, Number 4

TRUDY W. BANTA

Soil to Support 1,000 Flowers

In April 1993, I was in the audience at the annual meeting of the American Educational Research Association in Atlanta when Marcia Mentkowski, Donna Sundre, Glen Rogers, and Gary Pike gave their symposium on the topic "How Educational Research Can Contribute to Assessment in Higher Education." The more I listened, the more convinced I became that this was a set of papers *Assessment Update* readers needed to see. As soon as the session ended, I approached Marcia, an *Assessment Update* editorial board member, to ask, "Would you and your colleagues be willing to contribute your work for a special issue of *Assessment Update*?"

Marcia, Donna, Glen, and Gary generously agreed to adapt their papers for *Assessment Update*, and Marcia said that she would coordinate the project. Anyone who has attempted to rework an academic paper for a publication like *Assessment Update* will appreciate the effort involved in fulfilling that commitment. I express here, on behalf of all *Assessment Update* readers, immense gratitude to Marcia for her careful shepherding of the process and to Donna, Glen, and Gary for taking the time to develop their good ideas and observations into such cogent forms.

As C. Robert Pace observed in his Foreword to *Making a Difference: Outcomes of a Decade of Assessment in Higher Education* (Banta and Associates, Jossey-Bass, 1993), the fledgling discipline of outcomes assessment in higher education has brought back the 1930s' "virtues of variety and creativity" in measurement, evaluation, and educational research. I see many other indications that assessment practice is increasingly becoming rooted in the fertile soil of research in these and related fields. This growth seems enormously important since, to rephrase Santayana's well-known observation about history, those of us who ignore the wisdom available to us in these established fields are condemned to rediscover that wisdom on our own.

My experience in working with the 45 authors who contributed to *Making a Difference* affirmed my impression that, to date, most of the individuals

The fledgling discipline of outcomes assessment in higher education has brought back the 1930s' "virtues of variety and creativity" in measurement, evaluation, and educational research.

Assessment Update
January-February 1994
Volume 6, Number 1

selected to coordinate campus assessment programs have not been trained in measurement, education, or developmental psychology. Instead, they are faculty with expertise in political science, English, business, natural sciences, chosen to utilize their considerable leadership and communication skills to rally their colleagues around the difficult cause of assessment. However, there are signs that this situation may be changing. As more and more faculties elect to develop their own measures for assessment purposes, there is increasing recognition that they need access to colleagues with measurement skills. At the 1993 Assessment Conference in Indianapolis, an annual meeting that typically attracts an audience relatively new to assessment, I asked the assessment coordinators to identify themselves. Then I asked how many had backgrounds in education or psychology, and virtually all said that they did. These new coordinators will help to bring educational research and assessment into closer alignment.

I see creative work in assessment beginning to flourish as a result of building on research findings from several core disciplines.

I also learned from my coauthors in *Making a Difference* that faculty who take assessment seriously—who conscientiously develop multiple strategies for assessing student performance and then study the results to see what implications they may have for improving curriculum and instruction—are often led to consult the specialized literature of higher education to help them make meaning of the findings. I believe, for instance, that involvement in assessment has drawn more faculty from political science, English, business, and natural sciences into becoming familiar with what has been written about good practice in promoting student learning in college than would ever have been the case without assessment as a catalyst. Faculty assessors have discovered Chickering and Gamson's "Seven Principles for Good Practice in Undergraduate Education" (*Wingspread Journal*, 1987, 9[2], 1-4), Kuh, Schuh, Whitt, and Associates' *Involving Colleges* (Jossey-Bass, 1991), and Pascarella and Terenzini's *How College Affects Students* (Jossey-Bass, 1991). From these and other publications they have learned the importance of providing an environment in which students feel valued, setting high expectations for students' learning, regularly informing them about their progress, encouraging them to learn collaboratively by working in teams, and offering a variety of learning opportunities in order to accommodate a variety of student learning styles.

As another illustration of the ways in which I see creative work in assessment beginning to flourish as a result of building on research findings from several core disciplines, Peter Ewell and Dennis Jones of the National Center for Higher Education Management Systems (NCHEMS) have completed *A Preliminary Study of the Feasibility and Utility for National Policy of Instructional "Good Practice" Indicators in Undergraduate Education* (NCHEMS, 1993) for the National Center for Education Statistics. Recognizing that direct assessment of student learning has proved to be more difficult, expensive, and imprecise than early advocates had hoped it would be, Ewell and Jones suggest that we use as proxy measures indirect indicators of student learning. That is, building on the strong linkages that research has revealed between increases in student learning and such practices as promoting active learning, time on task, and involvement in the classroom, we could use student and faculty questionnaires and classroom

Assessment Update
January-February 1994
Volume 6, Number 1

observations to assess the frequency and quality of these practices, thus inexpensively and reliably assessing the conditions that support student learning. The very act of assessing these educational processes would lead to the improvements that faculty seek from more difficult-to-design assessment of student outcomes.

In the July-August issue of *Assessment Update* (1993, vol. 5, no. 4, p. 3) I wrote in this column about the request for proposals that the National Center for Education Statistics (NCES) had issued to initiate a national performance-based assessment system for college graduates. In late 1993, funding limitations imposed by Congress and other priorities within the U.S. Department of Education led NCES to withdraw this request for proposals. NCES officials nevertheless have expressed interest in trying out some less expensive assessment options in the short term. I, for one, would like to see NCES test the strategy that Ewell and Jones have proposed because it encourages in faculty and students the teaching and learning behaviors that research has demonstrated to be most effective in nourishing student development.

Assessment Update
January-February 1994
Volume 6, Number 1

Can We Combine Peer Review and National Assessment?

Can NCES help states assess the student learning described in Goal 6.5?

In early December 1995 I took part in a meeting in Washington that might be considered the end of an era. Or was it the beginning of another? We won't know the answer until the full impact of congressional intentions to make block grants to the states is apparent.

Sal Corrallo, senior staff member at the National Center for Education Statistics (NCES), and project director of the National Assessment of College Student Learning, called the meeting. *Assessment Update* readers will recall that Corrallo has led the effort to identify ways to measure college graduates' abilities to think critically, solve problems, and communicate effectively, the skills associated with National Education Goal 5.5, now 6.5 (see Corrallo's article in *Assessment Update*, Vol. 6, No. 3, pp. 5, 7, 11). Corrallo plans to retire in March 1996, and the result of the conference in Washington was to close the door on the federal effort he has shepherded and to test the feasibility of opening a door on future state efforts in this area.

The December conference was designed to acknowledge what NCES has been able to accomplish to date in connection with Goal 6.5 and to answer the question, Can NCES help states assess the student learning described in Goal 6.5? Those invited to attend the conference and join in this conversation were higher education executive officers in the 50 states, their representatives, or both, as well as a sprinkling of those of us with long-time interests in implementing postsecondary assessment. Joe M. Steele, one of the conference presenters, has prepared a summary of his remarks as the lead article in *Assessment Update* (Vol. 8, No. 2).

The NCES role has been quiet, but significant. In 1991, 15 of us were commissioned to write papers describing some of the issues surrounding national assessment of college student learning. Our papers were reviewed by an additional 30 individuals, and the 45 of us gathered in Washington in November 1991 for an initial discussion. A similar sequence of activities took place in 1992 in connection with a second set of commissioned papers.

Assessment Update
March-April 1996
Volume 8, Number 2

BEST COPY AVAILABLE

Following the second meeting, the National Center for Postsecondary Teaching, Learning, and Assessment at Penn State received NCES funding to conduct a series of Delphi studies designed to characterize the learning outcomes associated with the Goal 6.5 abilities of thinking critically, communicating, and solving problems. In addition, Peter Ewell and Dennis Jones of NCHEMS were commissioned to identify indicators of good practice that might constitute indirect measures of the three abilities. Both of these research projects provided evidence that it is possible to achieve agreement among postsecondary educators concerning the definitions of the complex skills sought by the National Education Goals Panel.

Hoping to build on the several activities initiated in 1991, Corrallo and others at NCES issued a request for proposals (RFP) in 1993 that, had the federal funding come through, would have produced a framework for assessing levels of achievement in each of the three Goal 6.5 areas as well as examples of assessment approaches. For 1996, NCES had planned to issue another RFP for the deployment of a national assessment of college student learning in the three areas. But in the aftermath of the 1994 congressional elections, the future of all National Education Goals activities is very much in question.

The December 1995 meeting produced a number of suggestions in connection with the basic question. In order to strengthen state assessment efforts, if that is to be the level at which the principal action will occur in higher education policy-making, NCES was encouraged to promote the development of state and regional networks of people interested in assessment by means of newsletters, listservs, World Wide Web sites, and state-supported pilot projects.

As I listened to the dialogue, another approach occurred to me. Sal Corrallo had asked me to report on Tennessee's experience with performance funding. As I have reported previously in this column, a study I conducted recently with Linda Rudolph, Janice Van Dyke, and Homer Fisher (*Assessment Update*, Vol. 6, No. 4, pp. 3, 14) suggested that peer review is the performance funding activity most productive of improvement on campuses in Tennessee. This includes program reviews conducted by individual campuses as well as disciplinary accreditation reviews. Over the years I have noted that most programs that are subject to disciplinary accreditation pay more attention to the need for assessment than those that are not guided by an accreditation body. Similarly, institutions tend to become most active in carrying out assessment just prior to the time of their ten-year regional reaccreditation process when that process emphasizes outcomes. I began to wonder if the answer to the NCES question about perpetuating its assessment emphasis might be to strengthen the assessment mechanisms utilized by disciplinary and regional accreditation bodies.

Most programs that are subject to disciplinary accreditation pay more attention to the need for assessment than those that are not guided by an accreditation body.

Federal influence on accreditation organizations in the realm of outcomes assessment is not new. In 1988 Secretary of Education William Bennett issued an executive order requiring these bodies to include assessment criteria in their guidelines if they were interested in recognition from his office.

Assessment Update
March-April 1996
Volume 8, Number 2

But despite the fact that most accreditors now have such criteria, not all have found effective ways to define and enforce them.

The public's interest in holding colleges and universities accountable for increasing student learning of critically needed citizenship and workplace skills is strong and growing stronger. Institutions must find ways of providing credible evidence of student learning that are acceptable to educators as well as to the public. Few in academe relish the thought of having either federal or state government officials tell them how student achievement will be assessed. The judgment of respected peers is the evaluation mechanism of choice for most academics. Thus, one sensible approach is to make the assessment of student outcomes as well as program and institutional effectiveness a central part of the peer review associated with disciplinary and regional accreditation. In this way we can accommodate the federal (and state) interest in assessing learning within the framework of peer judgment that is most acceptable to faculty and administrators.

Institutions must find ways of providing credible evidence of student learning that are acceptable to educators as well as to the public.

Although the national coordinating body known as the Council on Postsecondary Accreditation was disbanded in 1993, national leaders in higher education are working now to win institutional approval of plans for a new such organization, which would be called the Council for Higher Education Accreditation (CHEA). Leadership for CHEA would be provided by a board of directors composed of current chief executive officers of colleges and universities and trustees who are widely representative of the professions and the public at large.

For years colleges and universities have been providing information to NCES via surveys that form the foundation of the Integrated Postsecondary Education Data System (IPEDS). IPEDS surveys, which are completed by all providers of postsecondary education, currently are focused primarily on the collection of input data: student enrollments, faculty salaries, and institutional finances. But couldn't CHEA draw on its member institutions for the faculty experts needed to work with IPEDS in developing appropriate national indicators of student learning and institutional effectiveness? If so, NCES could utilize the national system that supports the peer review that academics value to collect the kinds of data about student and institutional outcomes that governments and the public find credible.

Assessment Update
March-April 1996
Volume 8, Number 2

Section V.

International Perspectives on Assessing Quality in Higher Education

I first noted the growing international interest in assessing the quality of higher education in my inaugural *Assessment Update* column in 1989 (1:1, which appears in the section on *Faculty Involvement*). At the first International Conference on Assessing Quality in Higher Education, which was held at Cambridge University in 1989, our participants were slowed in their efforts to communicate about assessment by differences in basic terminology, like the use of 'staff' in England to refer to 'faculty' and the use of 'course' to mean a course of study as opposed to an individual class. The following year (2:4) I wrote that participants in the Second International Conference held at St. Andrews University in Scotland, "armed with the glossaries we had prepared, took very little time to get acquainted and launched immediately into spirited discussions of government policies, faculty reactions, and methods of coping with calls for systematic evaluation of our work."

I finally felt that I had a truly global perspective on assessment.

In 1992 we published our first article about assessment at an institution abroad (4:3). That piece was written by an American who had spent some time at Portsmouth Polytechnic University in England. Since that time we have published about a dozen pieces by academics from England, Mexico, the Netherlands, South Africa, Hong Kong, and elsewhere.

By the end of 1992 I had visited China, Germany, Belgium, France, and the Netherlands, as well as England and Scotland, to talk with academic colleagues about assessment. In 4:6 I was able to speak with some confidence about quality assessment initiatives in these countries as well as in Austria, Sweden, and Australia.

By the time we had conducted the Sixth International Conference on Assessing Quality in Higher Education in Hong Kong in 1994, I wrote, "I finally felt that I had a truly global perspective on assessment" (6:6). In that column I developed a list of common concerns, like sharply rising costs and mobility of students and faculty across national boundaries, that were causing societies around the world to raise questions about the quality of higher education. I added concerns to that list in 7:2 and 9:6. In 7:2 I noted that "Faculty on both sides of the Atlantic seem to prefer evaluation by respected peers to any other method of assessment," but went on to note the varied palette of assessment methods developing worldwide. In 1997 (9:6) I wrote that while European participants in the Ninth International Conference held in Indianapolis tended to write in their papers about national assessment systems as opposed to the classroom and curricular evaluation approaches that were the subject of most U.S. papers, this distinction was beginning to blur. And I observed with satisfaction that, for the first time, almost all the feature articles in *Assessment Update* issue 9:6 were written by academics from countries other than the U.S.

Faculty on both sides of the Atlantic seem to prefer evaluation by respected peers to any other method of assessment.

TRUDY W. BANTA

Assessment: A Global Phenomenon

Peter Ewell and I had the extraordinary experience in the summer of 1990 of witnessing, in the space of just two weeks, the ways in which worldwide concerns about the quality of higher education are being played out in the West and in the East. In his column, Peter draws the outlines of the picture we saw in China. Here, I will focus on the stories told by our colleagues at the Second International Conference on Assessing Quality in Higher Education, held at St. Andrews University, Scotland.

The First International Conference had taken place in July 1989 at Cambridge University, England, with 45 individuals in attendance from nine countries. While we felt we had many common interests, we struggled to communicate—even though we all spoke English—because the terminology we use to describe our work differs (for instance, *course* in England refers to a program of study). The conference evaluation forms contained an urgent request: "Give us a glossary of terms!"

In July 1990, 80 participants representing 20 countries, armed with the glossaries we had prepared, took very little time to get acquainted and launched immediately into spirited discussions of government policies, faculty reactions, and methods of coping with calls for systematic evaluation of our work. Very recent developments in the United Kingdom and in Europe guarantee that interaction will be even more lively at the Third International Conference, scheduled for July 16-19, 1991, at Bath College of Higher Education, in England.

Ulrich Teichler, director of the Center for Research on Higher Education at the University of Kassel, told us that the quality of German universities has been assumed to be uniformly high: bright students, competent professors, and government research funds are (at least theoretically) distributed more or less equally among all German universities. In the last two years, however, the presumed cause-effect relationship between excellence in research and the production of well prepared graduates has been questioned by the federal minister of education and in an article published in *Der Spiegel*. The

Worldwide concerns about the quality of higher education are being played out in the West and in the East.

Assessment Update
Winter 1990
Volume 2, Number 4

latter revealed that German university students who responded to a nationwide survey gave higher ratings in some fields to new universities that lack distinguished research reputations. Teichler and his colleagues have undertaken studies designed to gather data from employers that will suggest ways to improve the academic preparation of graduates of German universities.

In most countries other than our own, as much as 90 percent of university funding is provided by the government. A theme we heard over and over at St. Andrews is that these governments are encouraging academics to establish self-study and peer-review mechanisms, which will produce qualitative judgments and recommendations that government officials can use in making policy and funding decisions.

In the United Kingdom, increasing use was made during the 1980s of peer review of the quality of research as a mechanism for determining funding levels. Now, Lewis Elton of the University of Surrey reports that the Committee of Vice-Chancellors and Principals has established two national units to help universities assess and improve the quality of teaching. The new Academic Audit Unit (AAU) will encourage each institution to define its own academic standards and build a system designed to achieve the delivery of these standards. Thus, each university will engage in its own process of self-validation, and AAU will undertake a critical review of that process. The new Universities' Staff Development and Training Unit will provide information and advice for universities seeking ways to improve teaching and learning.

Are there any skeptics in American higher education who still believe that the call for assessment, if ignored, will go away?

Don Westerheijden of the Center for Higher Education Policy Studies at the University of Twente, in the Netherlands, described the external evaluation system established during the last four years by the Association of Cooperating Universities in his country. Each disciplinary faculty conducts a self-evaluation. This is followed by a peer review carried out by a visiting committee, which evaluates all faculties within a given discipline in the 13-member association. Since 1988, the peer review system has been used to effect budget reductions, with the result that recommendations have been made to terminate or consolidate a small number of programs.

Felipe Rizo of the Universidad de Aguascalientes reported that in Mexico the National Association of Universities and Institutes of Higher Education has approved a plan for internal self-studies and external peer review by visiting committees, which will be used to evaluate university programs in eight fields. In 1992, these evaluations will become an important component of the process of determining institutional funding levels.

Do these stories sound familiar? Are there any skeptics in American higher education who still believe that the call for assessment, if ignored, will go away?

Assessment Update
Winter 1990
Volume 2, Number 4

TRUDY W. BANTA

Evaluation of Teaching and Learning in Germany

Terry Smith's article about assessment in British higher education is the first of what I hope will be a growing number of *Assessment Update* pieces about international quality assessment initiatives. It's a special pleasure to know that Terry has had the opportunity to work with Peter Findlay, because all of us who attended the First International Conference on Assessing Quality in Higher Education at Cambridge University in 1989 took away very positive perceptions of the work going on at Portsmouth Polytechnic. Now *Assessment Update* readers have a chance to learn from Terry's experience there.

In February, Peter Ewell and I, along with American colleagues Tom Angelo, director of the Academic Development Center at Boston College, and Jim Wilkinson, director of the Danforth Center for Teaching and Learning at Harvard, had the extraordinary experience of addressing an assemblage of German faculty leaders on the subject of evaluating teaching and learning. The meeting was organized by HIS (Hochschul-Information-System), sponsored by the Federal Ministry for Education and Science, and conducted in the beautifully appointed Wissenschaftszentrum (Science Center) in Bonn. Although the arrangements for this conference were not completed until December 1991, interest in evaluation in higher education is so intense in Germany that by February 20, 1992, the date of the opening session, 270 faculty, students, officials of state and federal government, and members of the press had registered.

What is behind this intense interest? First I offer a bit of background, distilled from an excellent paper by HIS associate, Edgar Frackmann, entitled "Quality Assurance in German Higher Education." By tradition, all German universities are considered to be of equal, and high, quality. Both teachers and students must meet exacting standards for entry to these institutions. Candidates for faculty positions must demonstrate their capacity for doing advanced research in their disciplines, and because they know their subject matter so well, they are presumed to be good teachers. Students must pass rigorous examinations to qualify for admission, and even with the doubling

By tradition, all German universities are considered to be of equal, and high, quality.

Assessment Update
May-June 1992
Volume 4, Number 3

BEST COPY AVAILABLE

of enrollments that has occurred in German higher education since the late 1970s, only 21% of the 19- to 26-year-old age cohort participates.

Since German professors are not promoted within their own institutions, but must move to another university in order to advance, it is assumed that talent in the professoriate is rather uniformly distributed among institutions nationally. Moreover, German students traditionally have moved freely from one institution to another in completing their studies. All of this commerce among institutions promotes the assumption that individual faculty members are responsible for the quality of their own work and that individual students are responsible for their own progress through the system. German universities have not assumed responsibility for nurturing students or promoting their progress.

In the 1980s, some questions arose concerning quality in German higher education. As plans for the European Community were developing, Germans began to look beyond their borders and to assess the competitiveness of their workforce. They discovered that it was taking German students longer to complete their studies—an average of seven years for the first degree—than was the case in neighboring countries. German graduates were more than 28 years of age before they entered the labor market, and decision-makers became concerned that German youth were spending their most creative and productive years engaged in education rather than in an occupation. Obviously, if this continued, it could impair German economic competitiveness.

They discovered that it was taking German students longer to complete their studies—an average of seven years for the first degree—than was the case in neighboring countries.

The assumption of equal quality among institutions was called into question when studies revealed differences among German universities in the time it took for students pursuing the same courses of study to graduate. Moreover, as the student population increased, more attended institutions close to home, and at present fewer than 25% of university students change institutions before completing their studies. As students increasingly identify with a single university, that institution must assume more responsibility for the education and progress of its own students.

Although selected by a given university, German professors are actually employed by the state, and during the 1980s several states began to call for institutions to take more responsibility for the progress of their students and for evaluating the quality of student outcomes. In 1990, a foundation supported by German industry that had previously provided support in the form of research grants began to award prizes to universities for reducing the duration of studies for students. In a crowning blow to the assumption of equal quality, *Der Spiegel* actually published a ranking of universities based on student ratings. Encouraged by states, industry, and the press, students themselves have begun to assume leadership for developing instruments and methods to assess the quality of instruction.

Assessment Update
May-June 1992
Volume 4, Number 3

The purpose of the meeting in Bonn in February was to promote interest in self-evaluation within German institutions of higher education. Examples were introduced from Great Britain and The Netherlands, where peer review models are prevalent, as well as from the United States. Despite the fact that evaluation of teaching is a relatively novel idea in Germany, and is apparently being greeted with the same—dare I say even *less?*—enthusiasm by German professors as by faculty in the United States, our audience was polite. Some even exhibited genuine enthusiasm when several of us began to describe the specific evaluative methods and tools that we use. I was particularly impressed by the active involvement, including poised commentary and questioning, of the students who were present. However, I also noted an almost total absence of women (only 7 of the 270 participants were females, and several of those were graduate students, not faculty members).

As Peter Ewell and I discovered in working with faculty in China, student evaluation of instruction is the current focus of the attention being given to evaluation in Germany. Using student outcome assessment to evaluate programs, as it has been developing during the past decade in our country, is not yet fully understood or appreciated in either country. Peter, Tom, Jim, and I left Bonn exhilarated by the knowledge that we had participated in an historic first step toward that understanding in Germany. We know we will not be the last Americans invited to participate in the process of faculty development on this topic in that important country.

I *was particularly impressed by the active involvement of the students who were present.*

Assessing Quality—Toward Broader Perspectives

I hope to be in a position . . . to provide strong links among planning, process, assessment, and improvement actions.

I am awaiting a flight that will take me from the annual meeting of the European Association for Institutional Research (EAIR) in Brussels, Belgium, to my new home in Indianapolis, Indiana. I have said goodbye to the wonderful people at the Center for Assessment Research and Development at the University of Tennessee, Knoxville, with whom I worked for the better part of a decade. They included Margery Bensey, who functioned so effectively as managing editor of *Assessment Update* since we started the publication in 1989. Many of the ideas that have turned piles of seemingly unrelated manuscripts into more orderly collections of readable articles have been Margery's. Thanks in large part to her diligent efforts, we never missed a deadline for sending copy to Jossey-Bass, and the editors in San Francisco told us that the material we sent is as close to being ready to publish as any they received. I would like to take this opportunity to thank Margery for focusing her considerable talents as an editor on *Assessment Update* for the past four years.

And now on to Indianapolis, where there is such excitement! Chancellor Jerry Bepko assembled an administrative team for Indiana University-Purdue University, Indianapolis, that hopes to develop new, more effective ways to provide education in an urban setting in the twenty-first century. The wonderful leadership, plus what seems to me to be an extraordinarily healthy interest on the part of both faculty and administrators in evaluating the effectiveness of their programs, drew me there to the new post of vice chancellor for planning and institutional improvement. I hope to be in a position to do what I have always known needs to be done—that is, to provide strong links among planning, process, assessment, and improvement actions. This is an ambitious agenda, but if ever there were a time and place where it might be accomplished, I believe it is in Indianapolis today.

I brought lots of luggage to Indianapolis: the editorship of *Assessment Update* and of a new book—*Are We Making a Difference?*, a work with fifty authors published by Jossey-Bass in 1993—and the planning of a seventh national conference, the Assessment Workshops at Indianapolis,

Assessment Update
November-December 1992
Volume 4, Number 6

BEST COPY AVAILABLE

in November, and of a fifth international conference, to be held July 19-21, 1993, in Bonn. I am particularly glad to find at my new institution strong support for international activities, because I believe it is increasingly important for higher education institutions around the world to share ideas and experiences.

For the foreseeable future, institutions of higher education in virtually every country will have to compete with other services designed to contribute to the welfare of the populace. None of us can escape the need to prove ourselves accountable, as do other stewards of funds both public and private. Moreover, since the best we can hope for is continuation of current funding levels, we must improve our effectiveness and offer a better education to more students at a lower cost. Outcomes assessment can help us do that.

Our fourth international conference on assessing quality in higher education, which took place in July in the Netherlands, drew one hundred participants presenting twenty countries. No national group was more vocal than the Australians, because recent government initiatives there aimed at linking university performance and funding have prompted educators to initiate dialogues with students, employers, and others about the meaning of quality in higher education. As is the case at most meetings these days, the interest group on total quality management drew such a crowd, from most of the countries represented, that it had to be subdivided to provide appropriate opportunities for discussion.

For the foreseeable future, institutions of higher education in virtually every country will have to compete with other services designed to contribute to the welfare of the populace.

In September, at the biennial fall meeting in Paris of the Programme on Institutional Management in Higher Education of the Organization for Economic Cooperation and Development, I heard more about changes within the Swedish education ministry that will prompt greater interest in self-evaluation on the part of Swedish universities. Since 1988, prompted by government action a coalition of Dutch universities experimented with national peer review of disciplinary programs of study. Also in the Netherlands, the Center for Higher Education Policy Studies at the University of Twente conducted very important systematic studies of the impact of policies designed to promote quality in higher education.

At a national meeting held in Bonn in February (see *Assessment Update*, 1992, 4(3), 3-4), several of us from the United States talked with students and employers who argued that the teaching skills of German professors should be evaluated. At the EAIR meeting in Brussels, an Austrian government official told me that the heretofore unquestioned competence of the professoriate in his country is undergoing a similar challenge.

Of the countries I've studied this year, it is England whose universities come closest to those in the United States in emphasizing student outcomes in their evaluations (see Terry Smith's article in *Assessment Update*, 1992, 4(3), 1, 2, 4). Portsmouth and Birmingham Polytechnics seemed to be doing the most interesting work in assessing student opinion on the subject of quality, among other indicators.

Assessment Update
November-December 1992
Volume 4, Number 6

A Global Perspective—at Last

Worldwide there are common concerns about higher education quality.

Assessment Update
November-December 1994
Volume 6, Number 6

Our Sixth International Conference on Assessing Quality in Higher Education, which was held in Hong Kong in July 1994, attracted participants from 28 countries around the world. This was our largest and most diverse audience to date, and for the first time we had significant representation from countries in the Far East in addition to Americans and Europeans. After six years of attending international conferences—all in the United Kingdom and Europe—as I listened to the featured speakers and concurrent session presenters in Hong Kong, I finally felt that I had a truly global perspective on assessment.

Worldwide there are common concerns about higher education quality stemming from similar cultural, economic, and governmental developments. Most nations are currently trying to provide higher education for the largest percentages of their populations ever to attend a college or university. Moreover, the ethnic and cultural diversity of those seeking higher education poses pedagogical challenges for the professoriate unlike any encountered previously. The costs of higher education are rising at a rate outstripped only by health care, and the resources that might once have found their way into institutions' budgets now must be spread over an increasing variety of services.

Mobility of students and faculty across national boundaries is a growing phenomenon. This prompts consideration of international standards that permit easy transfer of credits and credentials.

This is an age of consumerism. Students, parents, taxpayers, and donors want to know what they are receiving for their investment in higher education. They also want to know that education is taking place as effectively and efficiently as possible.

All of these international developments that are raising concerns about quality result in calls for assessment. And while the emphasis given to any particular approach may differ, the arrays of assessment strategies now

BEST COPY AVAILABLE

employed in countries around the globe are becoming more similar by the day. The comprehensive philosophy of Total Quality Management (TQM), with its strong measurement component, has spread like wildfire among the colleges and universities of the world. At the Hong Kong conference, papers on TQM were presented by representatives from Australia, China, the Netherlands, South Africa, Taiwan, and the United Kingdom, as well as the United States.

In some colleges and universities where a good deal of attention is paid to the needs of industry, ISO 9000 quality audits are being conducted. This approach is used in England, New Zealand, and Singapore, for example.

Peer review has long been used by accreditation agencies and many universities in this country. More recently, peer review has become a national strategy for ensuring higher education quality in The Netherlands, Finland, and the United Kingdom. Peer review is also being considered in Japan, where universities are only beginning to come under governmental scrutiny.

Australia and Scotland are experimenting with performance funding. Chile has joined the countries of Europe in the search for national indicators of performance.

Although many of the conference presentations in Hong Kong had as their focus national quality-assurance initiatives designed to serve accountability purposes, I find far more interesting the campus-specific assessment approaches designed for improvement purposes that are being carried out in the United Kingdom and in this country. While it seems to me that the national resources available for faculty development in the United Kingdom must be the best in the world, the U.S. capacity for institutional research, including assessment, is enviable. These strengths, combined with many other cultural and economic factors in the United Kingdom and the United States, have produced the world's most advanced, specialized, and improvement-oriented assessment practice.

The arrays of assessment strategies now employed in countries around the globe are becoming more similar by the day.

I am currently immersed in reviewing nearly 200 institutional case studies in assessment for a new book, the contents of which will be discussed in future issues of *Assessment Update*. My observation above about specialized, improvement-oriented assessment practice in the United States is based not only on reading and hearing the Hong Kong conference papers but also on the work with cases for the book. What I am seeing is an increasing number of creative evaluation studies undertaken by faculty who employ their research skills to conduct assessment in their own disciplines. The result is specific, detailed information about gaps in student learning, deficits in fundamental skills, and problems with the student services provided on campuses. This is the best information yet gathered for directing improvements.

Some of the articles in Vol. 6, No. 6 of *Assessment Update* illustrate my points. In the lead article, Patricia D. Murphy describes one institution's answer to the often-asked question, "How do we assess student outcomes in graduate programs?" Very simply, faculty in various disciplines are

Assessment Update
November-December 1994
Volume 6, Number 6

beginning to state desired learning outcomes more explicitly, then rate the extent to which they see those outcomes attained in course assignments, comprehensive written and oral exams, and theses—all the traditional methods of assessing the individual accomplishments of graduate students. The difference now is that in addition to a pass/fail notation or letter grade on an individual transcript, graduate faculty have specific data that can be aggregated across students and compared with their stated goals to ascertain where students are most successful and where some new method or methods might be needed to remedy a deficiency in learning.

The tool used by graduate faculty at North Dakota State University to add new purpose to the assessment of student learning is described by Murphy as "a simple rating scale" attached to each desired learning outcome. In their article, Peter A. Facione and Noreen C. Facione tell us about one such simple rating scale that they have devised to assess critical thinking ability as manifested in students' essays, course assignments, projects, and performances. If faculty can use this scale to assess something as complex and difficult to define as critical thinking, and graduate faculty can apply similar scales in assessing the advanced attainments of Ph.D. students, we may be witnessing a giant leap forward in assessment technology—developed not by measurement specialists but by faculty who are employing their research skills to solve assessment problems in their own disciplines.

We may be witnessing a giant leap forward in assessment technology.

Serendipity has brought together in this issue one study that resulted in the definition of a problem and another that suggests a potential solution to that problem. We already had a full issue that included Anita Gandolfo's article on why students do not go to class when the attendance policy article by William E. McMullin and Jon M. Young arrived in the mail. We made room for the two to appear together. Both articles are based on carefully designed evaluation research conducted by faculty who employed their disciplinary investigative expertise.

William S. Moore, guest Community College Strategies columnist, describes the Student Voices Project, a collaborative research project involving six institutions in the state of Washington in a study designed to enhance "understanding of how students evaluate their learning experiences." This is yet another example of faculty applying their research skills to create innovative approaches to assessment. And Moore promises more of this. He says that the Student Voices Project should "provide colleges with a model for further qualitative studies on other student populations."

In my last column, I promised something in each issue for newcomers to assessment. Peter J. Gray's Campus Profiles column describes Chicago State University's first initiative in assessment. In addition, William J. Struhar's article is an entertaining description of the process of preparing to assess student outcomes in general education. (He does think it's possible to tame that beast!)

Assessment Update
November-December 1994
Volume 6, Number 6

TRUDY W. BANTA

The Many Faces of Assessment Methodology

Last summer (1994) I was surprised to receive a letter from Frans A. van Vught postmarked Boulder, Colorado. I learned that Frans, director of the Center for Higher Education Policy Studies in the Netherlands, was visiting Peter Ewell and Dennis Jones at the National Center for Higher Education Management Systems in Boulder and had been introduced there to *Assessment Update*. In his letter Frans expressed interest in preparing an article for *Assessment Update*. I was delighted at the prospect and encouraged him to prepare the piece that is the feature article in this issue (Vol. 7, No. 2). Frans and his colleague Don F. Westerheijden have used the *Assessment Update* format well, accomplishing the amazing feat of characterizing recent quality-assessment initiatives in 14 countries in just over 2,000 words.

As Westerheijden and van Vught tick off the factors causing concern about the quality of higher education in Western Europe, the U.S. reader will have a strong sense of déjà vu. The factors are the same on both sides of the Atlantic: concern about the growth of public expenditures in general and the rapidly rising costs of higher education in particular; questions about the value of higher education in comparison with its cost; growth in student enrollments and proliferation of new fields of study, both of which contribute to the higher price tag; and internationalization of the work force, increasing the need for some standardization of expected educational outcomes across national boundaries.

Westerheijden and van Vught suggest that there is an unprecedented convergence of Western governments with respect to the extent of autonomy accorded higher education institutions, particularly with respect to quality assurance. On the Continent, where state control has been strong traditionally, institutions are being given more autonomy, but in exchange for demonstrating their quality in ways not required previously. In the United Kingdom, as in the United States, where state control has not been strong in the past, state regulation is increasing, and the need for institutions to provide concrete evidence of their quality has never been greater.

> ***On the Continent, where state control has been strong traditionally, institutions are being given more autonomy.***

Assessment Update
March-April 1995
Volume 7, Number 2

BEST COPY AVAILABLE

Note also the similarities between the methods of assessment being used in Western Europe and in this country. All nations have a strong interest in finding better ways to assess the quality of teaching. Countries on the Continent, such as the Netherlands, Norway, and Denmark, are experimenting with the kinds of voluntary self-study and external peer review of institutions that for many years have characterized regional accreditation processes in the United States. Disciplinary reviews carried out in France, the Netherlands, and elsewhere in Europe are similar to the kinds of program evaluation carried out by professional accreditation associations and in campus-initiated departmental peer reviews in this country.

Faculty on both sides of the Atlantic seem to prefer evaluation by respected peers to any other method of assessment. Westerheijden and van Vught point out that the disciplinary reviews conducted by the French Comité National d'Évaluation "are perceived by faculty as the most useful and valid aspects of the quality-assessment process." In a recent study of performance funding in Tennessee, Janice Van Dyke, Linda Rudolph, Homer Fisher, and I found that faculty considered peer review of undergraduate and graduate programs and disciplinary accreditation reviews to be the most effective of ten performance indicators in assessing the quality of higher education (*Assessment Update*, Vol. 6, No. 4, pp. 3, 14).

Faculty on both sides of the Atlantic seem to prefer evaluation by respected peers to any other method of assessment.

I am increasingly intrigued by the many methods of assessment that are being applied around the world. Each issue of *Assessment Update* is a sampler, and this one provides a diverse array of methods being applied on campuses across this country.

To look at interactions between students and campus services and among the units providing the services, Samford University in Alabama uses the quality-improvement strategy of service mapping. In charting the enrollment process, Don Belcher and Billy J. Strickland report that the Samford faculty and administration found that they could use number of applicants and number enrolled, number requesting financial aid and number receiving aid, number of students persisting to graduation, and number of graduates placed in jobs or graduate programs as measures for judging the effectiveness of the process.

Anita Gandolfo tells us that after developing a conceptual framework for assessment at West Virginia University, the faculty there became involved in a formative approach designed to improve teaching and learning. This took the form of a longitudinal study of student development wherein faculty interviewed students periodically to see how they were experiencing college.

Sandra Z. Keith urges mathematics faculty to become involved in campus efforts to assess student achievement of general education goals. In so doing, math faculty may be able to convince colleagues that quantitative literacy is best taught across the curriculum.

Assessment Update
March-April 1995
Volume 7, Number 2

Western Carolina University is the institution featured in Peter J. Gray's column (page 235). A team of faculty and administrators there designed an assessment resource guide and an approach to general education assessment that utilizes focus groups of faculty who teach general education courses; testing of freshmen and rising juniors; satisfaction surveys for students, faculty, and graduates; and employer follow-up.

In his column (page 168), Jeffrey A. Seybert describes methods for assessing the quality of noncredit continuing education programs. The methods he suggests include using the reoffering of a noncredit course as one measure of effectiveness and the migration of noncredit students into credit courses and programs as another. Information about licensure rates and about the extent to which students or participants achieve personal goals in a course or program are other indicators of noncredit program effectiveness.

Given my increasing interest in assessment methods, due in part to the experience of reviewing materials for my forthcoming book, we are inaugurating a new continuing column this month, *Campus Strategies*. In this column, the spotlight will be on particular strategies or methods for assessment being used on campuses today, reported by those who developed and/or implemented the methods. Each story will appear in a common format, using the following major headings: Background and Purpose, Method, Findings, Use of Findings, and Success Factors. Thus, the method will be described not only in terms of the way it was carried out but also in terms of the data it yielded and the changes that occurred as a result of paying attention to the meaning of the assessment findings.

The case described in the inaugural *Campus Strategies* in Vol. 7, No. 2 comes from Sharon S. Gibson-Groshon and Stuart Miller at Towson State University in Maryland. The method is use of a scoring guide for assessing student performance on measures of general education outcomes. One group of faculty used an inductive approach to develop the scoring guide while another group used a deductive approach. The two approaches were mutually validating and enabled faculty to use their own language in the criteria for judging students' work. This process of codifying faculty judgments as assessment criteria will be described further in future *Campus Strategies* columns.

Further Steps Toward Globalizing Assessment

> ***QA** both begins and ends at the classroom level.*

When Alan F. Howe gave the presentation on which his article in Vol. 9, No. 6 of *Assessment Update* is based, I was in the audience. Howe and I were among the representatives of 24 countries who were present at the Ninth International Conference on Assessing Quality in Higher Education, held July 21-23, 1997, in Indianapolis. While I have written in this space previously (*Assessment Update*, 1994, Vol. 6, No. 4, pp. 3, 13) about the similarities among nations in terms of the external pressures that are impelling institutions of higher education to adopt systematic approaches to assessing quality and the responses of faculty and administrators to these pressures, it seems that the similarities are multiplying and strengthening by the year. In countries as diverse as Canada and Brazil, Denmark and Greece, Russia and Singapore, South Africa and Thailand, stakeholders in higher education are concerned about delivering value for investment, accommodating the increasing numbers and diversity of students, covering cost increases when funding sources are becoming more limited, developing a work force with skills that are competitive in a global marketplace, and producing graduates with credentials that are transferable across cultural and national boundaries.

The responses to these concerns by higher education leaders worldwide have also contained similar themes. Elsewhere (*Quality in Higher Education*, 1997, Vol. 6, pp. 9, 3), I have used the analogy of concentric circles to describe the variety of responsive quality assurance (QA) systems represented in the conference papers at Indianapolis. In the outermost circle are national systems, followed by regional and state systems, then campus-specific institutional effectiveness programs, assessment in the disciplines, and, at the center, classroom assessment. QA both begins and ends at the classroom level. Reflecting the historical role of central government in higher education in most countries outside the United States, institutional leaders from the other countries represented at the Indianapolis conference were focusing their attention primarily on national and institutional QA systems. Most of the U.S. papers were about classroom assessment or assessment in the discipline.

Assessment Update
November-December 1997
Volume 9, Number 6

BEST COPY AVAILABLE

The three articles in this issue that describe QA in other countries reflect the differences in focus that I observed in Indianapolis. Laetus O. K. Lategan offers thoughts about the importance of self-assessment at the institutional level even as the South African Committee of University Principals is establishing a national Quality Promotion Unit to conduct institutional audits and promote "accreditation for purposes of articulation." Alan F. Howe and Chris Rust tell us how their universities have responded to the national QA initiatives of the Higher Education Funding Council of England that have been under way since 1992. I have long admired the many creative approaches to faculty development that can be found in British universities, and Howe and Rust do not disappoint in this regard. Look for Howe's description of the activities of the new Teaching Enhancement Office that he directs. And enjoy Rust's discussion of the work being conducted at Oxford Brookes University on "criterion-referenced marking-grids"—the British attempt to make grades more meaningful. For an American version of this same effort, please see an earlier article by Barbara Walvoord and Virginia Anderson (*Assessment Update*, 1995, Vol. 7, No. 6, pp. 8-9).

Society's interest in evidence of higher education's accountability for the resources it consumes has fostered shifts in the autonomy given to institutions that have served to bring colleges and universities around the world closer in yet another sense. In countries like England, where authority and funding for universities have historically come primarily through grants emanating from government bodies, the funding councils in the last decade have begun to give institutions more autonomy in exchange for increased evidence of accountability. In the United States, where higher education has been largely self-regulatory and independent of government controls, the past decade has seen strong action by states and even the federal government to influence both policy and practice in public colleges and universities and to exact much more in the way of accountability reporting. As a consequence, higher education in the United States has moved closer to the model of government control at work in other countries than has ever before been the case.

Higher education in the United States has moved closer to the model of government control at work in other countries than has ever before been the case.

In the early 1990s, the adoption of the National Education Goals by Congress brought U.S. higher education as close as it has ever been to having externally imposed national standards for college student learning. *Assessment Update* has attempted to keep its readers apprised of related developments in articles by Sal Corallo of the National Center on Education Statistics (NCES) (see *Assessment Update*, 1993, Vol. 5, No. 3, pp. 5, 7, 11; 1996, Vol. 8, No. 4, pp. 1-2, 15) and Elizabeth A. Jones, now at West Virginia University (see *Assessment Update*, 1993, Vol. 5, No. 6, pp. 8-9; 1996, Vol. 8, No. 6, pp. 7-8, 15).

With the election of the 1994 Congress, funds for developing a national system for assessing college student learning were withdrawn. But the National Center on Postsecondary Teaching, Learning, and Assessment at Pennsylvania State University did receive support to continue a series of national panel studies that were designed to explicate the meaning of the fundamental skills of writing, speaking, listening, reading, and problem

Assessment Update
November-December 1997
Volume 9, Number 6

solving. Jones's 1996 article in *Assessment Update* presented findings related to the first three skills on this list. In this issue, we are pleased to present her report on the final study in the series. Publication of Jones's findings on problem solving and critical reading outcomes bring to a close the NCES-supported initiative to establish national standards for college student learning that took place during the 1990s. Only time will tell whether or not this initiative will resurface in the new century as one more step in the ever-increasing globalization of higher education.

Assessment Update
November-December 1997
Volume 9, Number 6

From the States

Peter T. Ewell

About the Author

Peter T. Ewell is a Senior Associate at the National Center for Higher Education Management Systems (NCHEMS). A member of the staff since 1981, Peter's work focuses on assessing institutional effectiveness and the outcomes of postsecondary education. Peter conducts both research and direct consulting assistance with institutions and state systems on collecting and using assessment information in planning, evaluation, and budgeting. Peter has written many books on the topic of improving undergraduate instruction through the assessment of student outcomes. Prior to joining NCHEMS, Peter was Coordinator for Long-Range Planning at Governors State University and he was on the faculty of the University of Chicago. A graduate of Haverford College, he received his Ph.D. in Political Science from Yale University in 1976.

Contents

Page

***From the States:* Ten Turbulent Years** 151

The Current Pattern of State Assessment Initiatives 155
New State System Assessment Initiatives 159
Florida's College-Level Academic Skills Testing (CLAST) Program 162
Assessment in Colorado 164
Assessment in South Dakota 167
Survey of State Assessment Initiatives 170
Mandated Assessment: A Look in the Cultural Mirror 172
A New Approach to Quality? 174
Assessment in Hard Times: A Tale of Two States 177
Regional Accreditation: The Uncertain Alternative 180
Developing Performance Indicators for Community Colleges: Evidence from Two States 182
Institutional Consortia: A Promising Alternative for State Assessment Policy . . 185
A New Federal Role in Assessing Outcomes 187
"Program Excellence" in Ohio: An Indirect Approach to Assessment 189
Performance Indicators: A New Round of Accountability 192
More Federal Follies 194
Performance Funding in Texas 197
Part H: The Shape of the "Camel's Nose" 200
A New Look at Accreditation 202
Performance Funding: New Variations on a Theme 204
Acting Out State Postsecondary Reviews 206
Assessment in Tennessee: A State Profile 209
So, Are They Really Going Away? 212
Assessment Up Front 214
On the MARC: The Efforts, Obstacles, and Successes of the Maryland Assessment Resource Center 216

Page

Speaking Virtually: Assessment and the Western Governors' New University . . 220
Putting It All on the Line: South Carolina's Performance Funding Initiative . . . 223
New Looks in Accreditation 226
Technology Facilitates National Data Collection Strategies 229
Implementing Performance Funding in Washington State:
Some New Takes on an Old Problem 232
Statewide Testing: The Sequel 236

PETER T. EWELL

From the States: Ten Turbulent Years

It is always a bit humbling to be confronted with one's earlier prose, especially when one has been rash enough to make predictions. Certainly, such is the case with this column. *From the States* began straightforwardly enough as an attempt to review developments in mandated assessment on a state-by-state basis during a period in which state governments appeared to be the only likely accountability players. Since then, as this column has recorded, the external forces shaping assessment have become both broader and more complex. An (at the time) unthinkable set of federal initiatives has waxed and waned. For better or worse, voluntary accreditation finally began fulfilling its promise as a major player in stimulating assessment. Above all, inexorable structural changes in the ways postsecondary education is being delivered—including rising rates of multi-institutional attendance and a veritable explosion in the use of distance-delivered technology-based instructional alternatives—are increasingly rendering competency-based assessment the only viable way for external authorities to account for what is going on. In short, a lot has happened in a relatively brief span of time.

A lot has happened in a relatively brief span of time.

The first half-dozen pieces in this collection reflect a kind of "golden age" of institution-centered state-based assessment mandates that began in 1985 and concluded in about 1992. Whether established by legislation-like Colorado's HB1187 and South Carolina's Act 255—or by a "preemptive strike" on the part of a state higher education agency to head off prospective legislation—as in states like Virginia or New York—the state initiatives chronicled here allowed institutions substantial flexibility in determining both what should be assessed and how to do it. Cautionary tales, like those of South Dakota and Washington (the latter reviewed in an excellent article by Bob Thorndike published in *Assessment Update*'s summer 1990 issue) pointed out the presumed folly of relying on single indicators and standardized tests to discharge accountability functions. At the same time, reviews of well-established test-based initiatives like Tennessee's Performance Funding initiative and Florida's "rising junior" examination program served largely to confirm widespread perceptions that such programs were special cases. The major exception here was New Jersey, whose higher education

BEST COPY AVAILABLE

agency launched what was easily the most proactive and technologically-sophisticated assessment effort of the decade. Reviewing this program in an early 1990 column, I wrote, "if assessment fails to meet its political objectives here, it is unlikely to succeed anywhere." For reasons partly noted in a later column ("A Tale of Two States" published two years later), it did not. More significantly, perhaps, the agency that conceived this initiative (the New Jersey Department of Higher Education)—together with nearly a dozen of its fellows—had itself become a casualty of increasing political turmoil.

Meanwhile, a number of national developments came to dominate the accountability scene. First, the early nineties witnessed the emergence of unprecedented federal initiatives related to assessment including "Student Right-to-Know," the National Education Goals process, and the "Part H" accountability provisions of the Higher Education Act (the last of these best known for its establishment of the State Postsecondary Review Entities, or "SPRE's"). Remarkable in retrospect was how quickly this federal accountability push developed and the incredible momentum that it rapidly acquired. Also worth noting was the fact that this was an international phenomenon: most other countries were also at the time experiencing strong government-led accountability mandates. Second, and in part stimulated by federal developments, voluntary accreditation finally emerged during this period as a major player, in the form of the short-lived National Policy Board for Postsecondary Accreditation (NPB). Both the SPRE's and the NPB went down quickly in the wake of the Congressional elections of 1994—though both have had some lasting effects. What the columns don't capture was the extreme urgency that many felt at the time, as a national accountability movement appeared inexorably to be emerging for American postsecondary education. In this atmosphere, most of us believed that almost anything could happen—and, for most of the wrong reasons, it turned out that we were right!

While performance funding remains overwhelmingly popular as a concept among state legislators . . . there is often insufficient political stability in the states adopting such initiatives to sustain the many delicate technical and operational negotiations needed to make them work.

During this same period, state attention moved decisively away from institution-centered assessment approaches toward performance indicators and performance funding. Partly this was a product of changing conditions—tight budgets and greater calls for "return on investment" from increasingly conservative taxpayers were now the norm. Partly, though, this policy shift represented the failure of "institutional improvement" approaches to assessment to effectively generate "bottom-line" data about performance of any use to government. This phenomenon is still under way—the latest and most complex instance being illustrated by one of the last columns in this collection describing South Carolina's performance funding effort. Again, what the several columns on these topics do not reflect is the extreme volatility of the current state political environment. While performance funding remains overwhelmingly popular as a concept among state legislators—and new initiatives are regularly launched as a result—there is often insufficient political stability in the states adopting such initiatives to sustain the many delicate technical and operational negotiations needed to make them work. Attempts at implementing 100% performance funding in Florida and Texas (the latter described in Jerry Gaither's guest column in December, 1993) eventually fell flat for these reasons. And as of this writing, both Arkansas and Kentucky are abandoning their performance funding approaches.

Another trail worth some reflection is the role of accreditation as an external force fostering assessment. When I labeled it the "uncertain alternative" in a 1992 column, we were in the middle of a period characterized by strong government resurgence. Under federal guidelines adopted during the reauthorization of the Higher Education Act of 1987, all regional accrediting bodies had by that time officially adopted some kind of assessment requirement. But few institutions felt much affected by accreditation in comparison to what the states were doing. Meanwhile, government players at both the state and federal levels were becoming increasingly skeptical about the ability of "insider" membership organizations like the regional accrediting bodies to discharge accountability functions credibly and meaningfully. Clearly, it was to higher education's collective advantage to support accreditation-based approaches to assessment in preference to state mandates outside the academy's control. But private institutions—unaffected by government mandates and constituting a powerful voice in accreditation circles—continued to resist external scrutiny in any form. As a result, accrediting bodies moved comparatively slowly in implementing their assessment provisions, further eroding political confidence in their quality-assurance capacity. Part H, and particularly SPRE, changed the political equation markedly in 1992-94, however. For the first time, powerful private institutions faced the threat of direct government review, and strengthening accreditation now looked pretty good in comparison. Although that particular threat subsided with the demise of SPRE in 1994-95, the "wake-up call" to accreditation was unmistakable. As a result, accreditors have in the ensuing period become far more innovative in experimenting with new approaches. While none of their initiatives have the force of the original NPB proposals, most regionals are at last beginning to fulfill the accountability promise advanced for non-governmental approaches to assessment a decade ago. And as noted by the penultimate column in this collection ("New Looks in Accreditation"), most all of them are engaged in significant process innovations.

We will probably see accountability policies—and with them assessment requirements—increasingly driven by new patterns of student attendance and new modes of instructional delivery.

In retrospect, what do the events chronicled by these columns add up to for the future? First, I think, they reinforce the conclusion that higher education accountability—and with it assessment—will always be mainly a state-based phenomenon. Given the frequent attention paid in these columns to federal, international, and associational activities, I've often been tempted to change the byline, *From the States* to something more comprehensive. Given the likely course of events in the next few years, it's probably wise to keep the focus as it is. Second, as reflected by the last column in the collection, we will probably see accountability policies—and with them assessment requirements—increasingly driven by new patterns of student attendance and new modes of instructional delivery. Technology-based instruction, multi-institutional paths of enrollment, and distance-based programming are here to stay and, whether we like it or not, they will profoundly alter the traditional structure of higher education. As the column on the emerging Western Governors' University makes clear, assessment is at the heart of these developments and will increasingly influence structures of accountability. Finally, despite what many of us might wish, external requirements will probably continue to drive what institutions do in the name of assessment far more than the imperative of internal improvement. As a result, it will be as important as ever for us to track what outsiders are thinking and what they are likely to do. I sincerely hope that *From the States* has helped in this regard and that it will continue to do so. Certainly, it has been fun to write!

External requirements will probably continue to drive what institutions do in the name of assessment far more than the imperative of internal improvement.

PETER T. EWELL

The Current Pattern of State Assessment Initiatives

I*s the character of state assessment initiatives changing?*

The purpose of this column is to periodically review the overall status and implications of state-mandated assessment initiatives. In each issue of *Assessment Update*, I intend to note any salient developments that have come to my attention and to provide a somewhat longer analytical description of one state's approach. I hope by this method to provide reasonable coverage of a complex and quickly changing topic.

In this initial issue, however, I would like to vary the promised program by quickly sketching the current pattern of state assessment initiatives. A first and reasonable question is how many there are. Obviously any count depends on how the question is framed. In its follow-up to "Time for Results," for example, the National Governors' Association reports only thirteen states without an assessment program either "in place or under consideration." ACE's "Campus Trends" survey for 1988 paints a different picture: 29 percent of institutional respondents report that their state "is currently requiring assessment procedures." Clearly, a count of these initiatives depends on definition. My own personal count of about fifteen states is based on the following criteria: (1) the existence of a formal document, enabling legislation, executive order, or board resolution on assessment; (2) an identifiable staff assignment or budget line for the program; and (3) a distinct and concrete requirement that institutions report progress or results. This count represents three more states than a year ago—a clear addition, but by no means the runaway train predicted by some. My best guess is that this rate of growth will remain steady for the foreseeable future—that is, another two or three states probably can be expected to mandate something each year.

Another question commonly posed is whether the character of state assessment initiatives is changing. Here the answer is definitely yes. First, in contrast to initial mandates that often had their origins in the state house or governor's mansion, the newer initiatives are coming from governing and coordinating boards. The past year saw no new legislation on assessment of the character of Colorado's HB1187 or Texas's SB543. Legislative

Assessment Update
Spring 1989
Volume 1, Number 1

action, if present at all, has occurred in response to board request. This trend indicates an important seizure of the initiative by system governing and coordinating boards and may significantly affect what is requested of institutions. On the one hand, institutions can probably expect assessment mandates to be better adapted to institutional realities; on the other hand, there also may be a greater tendency for assessment to be treated in familiar fashion by state agencies—as primarily a bureaucratic reporting requirement.

Finally, state mandates continue to emphasize institutional initiative. Their modal tendency is to require institutions to develop assessment plans covering such areas as basic skills and remediation, general education, achievement in the major field, student satisfaction, and alumni placement and achievement, with a fairly open-ended set of guidelines. Generally, institutions are required to report progress in implementing their plans on an annual basis. In addition, beginning in the second or third year of implementation, most guidelines require institutions to report results in summary form, together with a description of actions taken locally to improve effectiveness.

Institutions will continue to have many opportunities to help shape state requirements that have not been fully nailed down.

All told, the picture is one in which individual institutions retain considerable discretion—if they choose to exercise it. At the same time, institutions will continue to have many opportunities to help shape state requirements that have not been fully nailed down. But they are well advised to act quickly. As a consequence of seizing the initiative, state boards are under increasing pressure to show action. And without a credible institutional response, they may show considerably less flexibility than has been the case up to now.

State Notes

Kansas became the most recent state to mandate assessment through Board of Regents action adopted in April. The policy, typical of most current initiatives, requires each institution to describe its undergraduate program assessment plan to the Board by January 15. Progress reports were due in October. No specific instruments or approaches are specified, and institutions are "encouraged to utilize assessment procedures to improve program development and management strategies."

In September 1988, New Jersey completed the first phase of the pilot testing for its statewide General Intellectual Skills Examination scheduled for full implementation next fall. The examination, being developed under contract with ETS, is constructed around extended tasks up to ninety minutes in length, scored by faculty readers. The September pilot test took place in intact classrooms at eleven public and private institutions. A final pilot test took place in spring 1989.

Assessment Update
Spring 1989
Volume 1, Number 1

In Colorado, institutions took the first official step toward compliance with HB1187 by submitting reports on educational goals to the Colorado Commission on Higher Education last June 1988. These reports were required in advance of institutional assessment plans submitted in December

1988. First substantive reports of results and progress were required in October 1989. Under the legislation, all institutions submitted evidence of "demonstrable improvements in student knowledge, capacities, and skills between entrance and graduation" by 1990 or risked loss of up to 2 percent of their appropriation.

In Texas, full implementation of the Texas Academic Skills Program (TASP) was expected in fall 1989. Mandated by the legislature, the program calls for basic skills testing of all entering college freshmen, for mandatory placement and retest of those found deficient, and for reports on the effectiveness of remediation provided. Similar to the New Jersey basic skills program in place since 1977, examination instruments were designed by National Evaluation Systems (NES).

Assessment in Virginia: A State Profile

The origins of Virginia's assessment initiative date back to 1984 to the publication of NIE's report *Involvement in Learning*. After reading this report, several Virginia legislators were persuaded that some kind of assessment was needed to ensure quality. At the time, several states, most prominently Florida and Tennessee, had already mandated statewide testing, and Virginia legislators initially were inclined to follow suit. But they were quickly persuaded by the State Council for Higher Education in Virginia (SCHEV), the state's higher education coordinating body, to propose a study resolution instead. In 1986 the Council completed its study, and using it as a guide, the legislature mandated that each of the state's public higher education institutions develop its own plan to assess the outcomes of undergraduate education to be submitted to the Council for approval. Plans were to be submitted by June 1987 and were prepared within general guidelines established by the Council and a committee of institutional representatives. A strong inducement for institutions to cooperate was provided by a link between plan approval and access to state incentive funding: if an acceptable plan were not provided, the governor would not recommend institutions for such funds.

A strong inducement for institutions to cooperate was provided by a link between plan approval and access to state incentive funding.

By 1987, each institution in Virginia had an approved assessment plan that is now in the process of being implemented. True to the spirit of the program, Virginia institutions have developed a range of diverse, often innovative, assessment approaches, supported by a total of $4.4 million in state funds allocated to institutions for local assessment purposes over the 1988-1990 biennium. Beginning in July 1989, institutions will be required to report information within the categories established by the guidelines—outcomes in the major and general education, basic skills proficiency and the effectiveness of remediation, and alumni follow-up. In each of these categories, institutions are to present summaries of the results obtained, describe their reliability and validity, analyze what they mean, and describe any institutional action taken as a result.

Assessment Update
Spring 1989
Volume 1, Number 1

By counting heavily upon institutional initiative, Virginia's assessment approach has been called both exemplary and risky. Certainly it allows for considerable flexibility in the ways institutions can approach the task of building assessment. Emerging programs at James Madison University, George Mason University, Longwood College, the College of William and Mary, and many others are innovative and are tailored to the mission, clientele, and culture of each institution. Such programs provide evidence that, given proactive local leadership, such a program can work. But the Virginia approach is also risky. Banking heavily on institutional initiative first carries with it the substantial possibility of uneven response. Some institutions will be proactive, while others may do little or nothing. What to do about the latter remains a compelling problem for SCHEV. Second, there is a potential problem in using the resulting information for accountability purposes. Virginia's approach is premised on the proposition that accountability and academic improvement need not be in conflict and that programs that result in improvement also will serve to demonstrate effectiveness to the legislature and the public. Much, therefore, depends upon the State Council's ability in the coming year to translate a necessarily diverse and noncomparable set of institutional assessment reports into policy information that will satisfy the legislature and the public.

How Virginia fares may help determine the shape of assessment in many other states.

Because of these features, I believe that events in Virginia should be watched closely in the coming year. Many states now debating the best assessment approach will be interested in both the overall quality of programs built by Virginia institutions and the degree to which external constituencies in Virginia are satisfied with the information provided. How Virginia fares may help determine the shape of assessment in many other states.

For more information, Dr. Margaret Miller, assistant director for academic programs, State Council for Higher Education in Virginia, may be contacted.

Assessment Update
Spring 1989
Volume 1, Number 1

PETER T. EWELL

New State System Assessment Initiatives

Several new state system assessment initiatives are emerging, and those mandates already in place are becoming more fully institutionalized. Some of the most prominent recent developments have occurred in the following states.

Several new state system assessment initiatives are emerging.

Institutions of the State University of New York (SUNY) have begun developing local assessment programs for instructional improvement. A December 1988 memo from SUNY's central office in Albany directed all campuses in the system to submit, by June, a preliminary report that indicates what kinds of campus assessment activities are already in place. Over the 1989-90 year, institutions were expected to develop comprehensive multi-year assessment plans, due in June 1990.

In Virginia, a bill was passed that would make assessment a permanent responsibility, rather than merely a special assignment, of the State Council of Higher Education for Virginia (SCHEV). Proponents believed that passage of the bill would allow SCHEV to make a better case for incorporating funding for assessment ($4.5 million in the last biennium) into institutional base budgets. Opponents feared that passage would encourage greater state-level direction of a decentralized institutional process.

In Colorado, institutional accountability plans, prepared as a result of House Bill 1187 and submitted to the Colorado Commission on Higher Education, have been formally reviewed by an external agency and approved. While many plans involved standardized testing based on the usual array of instruments, others proposed some form of curriculum-embedded assessment based on faculty judgment. HB1187 authorized the commission to withhold up to 2% of any institution's appropriation if it is deemed out of compliance with the bill's accountability provisions.

In Florida, fears about low levels of anticipated minority performance on the state's CLAST "rising junior" examination prompted the commissioner of education to consider a delay in implementing higher pass-rate standards scheduled for 1989. A committee was asked to review the matter and at the

Assessment Update
Summer 1989
Volume 1, Number 2

BEST COPY AVAILABLE

same time to consider ways to make the state's extensive K-12 and postsecondary testing programs more coherent.

Assessment in New Jersey: A State Profile

New Jersey's College Outcomes Evaluation Program (COEP) is probably the most extensive, visible, and proactive initiative in the current new wave of state assessment mandates. Considerable resources, strong support from Governor Thomas Kean and the legislature, and active leadership from the state's Department of Higher Education (DHE) combine in New Jersey to provide an unprecedented range of activities. New Jersey's effort also has an unusually long time horizon. Formally launched in 1985, the program's complete implementation extends into the 1990s. New Jersey's program is both more systematic and more centralized than other state initiatives: multiple statewide task forces actively led and supported by DHE produced an extensive set of guidelines and procedures. At the same time, high levels of central coordination and support rendered conceivable the development of some common statewide measurement initiatives.

New Jersey's College Outcomes Evaluation Program (COEP) is probably the most extensive, visible, and proactive initiative in the current new wave of state assessment mandates.

In 1985, Chancellor T. Edward Hollander outlined a set of principles and guidelines for a program stressing the use of assessment information for instructional improvement. Two years of deliberation by a 25-member COEP advisory committee, consisting of faculty and administrators from across the state, shaped a program adopted in 1987. It included a common statewide general intellectual skills examination, based on a new technology of cognitive assessment, and a range of local evaluation activities, to be carried out by each institution. COEP's first component broke new ground in at least two ways. While a number of states—among them Florida, Georgia, South Dakota, and Tennessee—had previously developed common postsecondary testing programs, none had chosen the difficult route of building its own examination designed explicitly for comprehensive program evaluation. As a result, the coverage of the New Jersey examination is distinctive—an identifiable set of cross-cutting collegiate skills, described in terms of "gathering, analyzing, and presenting information." Second, the proposed examination, while standardized, is not based on such traditional recall-based methods as multiple choice. Instead, its core consists of a series of judgmentally scored tasks requiring students to comprehend, manipulate, and communicate complex bodies of information. To develop the required instrument, DHE contracted with the Educational Testing Service (ETS) to work with faculty teams in 1987-88. Initial tryouts of task-based items took place in the spring of 1988, and full-scale pilot tests were underway at approximately a dozen campuses throughout the 1988-89 academic year.

Under COEP's other components, institutions in New Jersey are developing their own procedures for assessing the effectiveness of instruction in general education and major fields, and of assessing students' personal development, satisfaction, and behavior after college. Local evaluation processes also address faculty, research, and creative activity, as well as the institution's impact on the community and society. Institutions submitted progress

Assessment Update
Summer 1989
Volume 1, Number 2

reports on their implementation of these activities in June 1988 and reported initial results in 1989. Rather than providing institutions with a simple set of assessment guidelines, as in Virginia or Colorado, New Jersey's approach gives campuses substantial advice about appropriate methodologies, in the form of specially prepared manuals, literature reviews, and accompanying workshops. These documents summarize the available techniques, approaches, and background literature related to a particular assessment area, in a form suitable for use by local practitioners. Information produced by local assessment is also supplemented by data drawn from available statewide databases.

Overall, COEP constitutes an instructive and distinctive field experiment of assessment practice that must overcome some formidable risks to be successful. Because of proactive leadership, state policymakers are incurring the risk that the effort will be perceived by institutions as a top-down initiative, lacking visible mechanisms for inducing institutional ownership and response. Technical risks involved in developing and fielding the general intellectual skills examination are also daunting. Because of its task-based approach, the instrument must overcome significant challenges with respect to scoring, the reliability of obtained results, and the great unknown of student motivation.

COEP represents a major test case for state-based assessment.

But the potential payoffs of New Jersey's assessment approach are also considerable. For New Jersey, COEP has the potential to provide policymakers with a comprehensive common database in terms of which to concretely demonstrate the state's return on investment in higher education. For the rest of us, COEP's pioneering task-based general intellectual skills examination represents an excellent alternative to large-scale multiple-choice testing. If it is successful, we will no doubt soon see versions of this technology marketed by ETS and others. Similarly, New Jersey's careful process of instrument and method review has produced a set of documents useful to any institution contemplating assessment; the COEP advisory committee's report alone is a valuable primer of assessment practice.

COEP represents a major test case for state-based assessment. A massive effort, clear political support, and a long timeline imply that if the initiative fails to achieve its political objectives here, it is unlikely to succeed anywhere else. All told, over the next few years New Jersey will be an interesting place to watch.

Florida's College-Level Academic Skills Testing (CLAST) Program

Rapid population growth, consequent increases in state budgets, and strong political support from both the legislature and the governor all contributed to a favorable climate for a proactive state role.

In Washington, the Higher Education Coordinating Board resolved to use institution-proposed outcomes measures as an alternative to standardized tests for sophomores. A $300,000 study mandated by the state's Master Plan for Higher Education concluded that the ACT-COMP, ACT-CAAP, and ETS Academic Profile were not suitable for state purposes.

In Illinois, institutions were required this summer to submit formal comprehensive reviews of undergraduate education to the Illinois Board of Higher Education. Review guidelines included explicit questions concerning student assessment. In contrast to other states, Illinois does not have a distinct assessment initiative, preferring to use regular statewide reporting and review mechanisms instead.

In Texas, the first students were tested for entering basic skills under the state's new TASP program, using instruments designed by National Evaluation Systems. Students cannot remain enrolled beyond the completion of 15 hours of regular coursework without having passed TASP.

Assessment in Florida: Review of the CLAST Testing Program

Probably the best known current example of statewide assessment is Florida's College-Level Academic Skills Testing (CLAST) program. The roots of CLAST date from 1973, when Florida began a major effort to upgrade its educational systems through direct state action. Rapid population growth, consequent increases in state budgets, and strong political support from both the legislature and the governor all contributed to a favorable climate for a proactive state role. CLAST's immediate impetus was a trial competency test of teachers in the St. Petersburg region, which obtained disappointing results. Reinforced by concerns about grade inflation and other issues in teacher education programs, the mood of the legislators and the governor was favorable for extending the state's growing K-12 competency testing movement into higher education.

Assessment Update
Fall 1989
Volume 1, Number 3

BEST COPY AVAILABLE

Initiated on a voluntary basis in 1979, with considerable faculty participation, the program's initial task was to define and assess a relatively narrow range of essential academic skills, including reading, writing, and computation. By 1982, an instrument was ready for general use. It was designed primarily at Florida State University under the guidance of numerous statewide committees, and the first official scores were to have been recorded in 1984. Beginning in that year, students would be substantially barred from junior-level coursework if they had not passed CLAST after three attempts.

Because of its high-stakes nature CLAST was immediately controversial and has remained so. Adverse minority impacts and widespread perceptions of "teaching to the test" are particular concerns. In both cases, attention has centered on community colleges; the vast majority of four-year college students can easily pass the examination at current standards. While both objections have some merit, the case is far from decisive. Overall pass rates on CLAST have averaged 95% after three tries over the last several years, with black and Hispanic pass rates at 87% and 92%, respectively. In general, student performance has kept pace with rising standards.

Many continue to question whether there is too much testing in Florida and . . . too much "micromanagement" of institutions.

Vocal opponents of CLAST, such as Miami-Dade Community College, have claimed significant adverse impacts as a result of testing (for example, substantial declines, particularly among minorities, among students enrolling in AA transfer degree programs). Other institutions, in contrast, have claimed positive impacts in the form of needed curricular realignment and more uniform standards. All institutions, moreover, have benefited from the additional funds appropriated by the legislature to implement CLAST reforms.

Most state policymakers currently see CLAST as having accompanied its original and rather narrowly defined objectives: consequently, they are not included to extend the program. Moreover, Florida's ten-year effort to reform higher education through direct state action now appears to be losing its momentum. A new and less supportive administration is in office, and level state budgets are expected over the next few years. For CLAST, these changes are most visible in a recent proposal to delay implementation of higher 1989 pass-rate standards. While the political issue of adverse minority impact is real, the main reason for considering delay was to rethink the state's strategy for assessment and reform. Although the 1989 standards were eventually implemented as planned, many continue to question whether there is too much testing in Florida and, by implication, too much "micromanagement" of institutions.

Despite current fashion, observers of assessment should be particularly cautious about generalizing the experience of CLAST to other states and should remember that Florida's testing program was intended to achieve very limited ends. Thoughtfully inspected, however, the CLAST experience can provide many lessons about how best to initiate and respond to a statewide initiative.

Assessment in Colorado

Colorado remains one of the few states where an assessment program was initiated by the legislature.

Assessment Update
Winter 1989
Volume 1, Number 4

In New Jersey, severe budget shortfalls have threatened full implementation of the state's ambitious College Outcomes Evaluation Program (COEP) for 1989-90. The legislature cut funds for the politically controversial statewide General Intellectual Skills test, but strong support from the governor has allowed some funding to be restored. The test was scheduled to be given in February.

In Florida, higher 1989 standards for the state's rising-junior CLAST examination will be implemented in English and reading, as planned. Math standards will be set midway between 1986 and 1989 levels, while essay scores will be held at 1986 levels for another year. Originally, the commissioner of education had recommended no increases in standards, largely because of adverse minority impact.

In Connecticut, guidelines for institutional assessment were issued in July, 1989 by the state's Department of Higher Education. Assessment was required in general education, major programs, basic skills and remediation, retention, student development, and alumni satisfaction. Final plans were submitted in June 1990, and institutions reported results biennially thereafter.

Assessment in Colorado: A State Profile

Colorado remains one of the few states where an assessment program was initiated by the legislature. In 1985, the Colorado General Assembly passed the Statewide Accountability Act (HB1187), requiring each of the state's 28 public institutions of higher education to be "held accountable for demonstrable improvements in student knowledge, capacities, and skills between entrance and graduation." In addition to citing cognitive development, the language of HB1187 noted additional dimensions of potential student development, including "self-confidence, persistence, leadership, empathy, social responsibility, and understanding of cultural and intellectual differences." To ensure compliance, the bill authorized the Colorado

Commission on Higher Education (CCHE), beginning in July 1990, to withhold up to 2% of the operating appropriation of any institution failing to implement the bill's provisions.

Despite HB1187's directive language and strong penalties, the essence of Colorado's assessment approach is decentralized. No statewide instruments are mandated, and institutions are given considerable discretion to define their own assessment approaches. In designing the bill, Senator Al Meiklejohn, its sponsor, was thinking explicitly about the accountability process required of Colorado's elementary and secondary schools—largely a public self-study, with considerable local community input and involvement. One element of Meiklejohn's intent was to stimulate local improvement in this way; another was to send an unambiguous signal to higher education that its responsiveness was also required.

From the beginning, CCHE encountered difficulties in developing guidelines for institutional plans. The fact that most assessment elements were written explicitly into the law severely limited CCHE's maneuvering room. HB1187 also redefined CCHE as a strong statewide coordinating board, authorizing it to establish mandatory institutional admissions requirements, to discontinue duplicative programs, and to establish a statewide funds-for-excellence program. Short of staff, CCHE had little choice but to implement such provisions piecemeal, with their priorities dictated by the target dates specified in law.

CCHE's new responsibilities also raised wider governance issues. Unlike the State Council in Virginia or the Department of Higher Education in New Jersey, which may negotiate directly with institutions on the development of assessment policy, CCHE has to work through the state's five separate governing boards. One result was further delay in implementation and, more important, impediments to the kind of direct communication necessary to prevent misunderstanding. Finally, no new funds were earmarked for assessment; institutions were expected to pay for their own programs with existing resources.

Many faculty were more opposed to the implication that the state could legitimately approve instructional goals than they were to assessment.

Institutional responses were required in two stages. In June 1988, institutions submitted statements of their goals and objectives for undergraduate education. Full-scale assessment plans were required the following December. While this sequence was conceptually sound, many faculty were more opposed to the implication that the state could legitimately approve instructional goals than they were to assessment. At the same time, the 2% budget penalty rendered plan approval a high-stakes process. As in Virginia, the state's strategy was to enhance credibility by involving external consultants in the review. Unlike what happened in Virginia, however, institutions were visibly ranked for comparison purposes. Seven of 28 submitted plans were not approved until July 1989.

As in most states that follow a decentralized approach, assessment plans in Colorado differ widely. Approximately half the plans propose administering at least one nationally normed standardized general education examination.

Assessment Update
Winter 1989
Volume 1, Number 4

Others concentrate on major-field assessment, usually embedded in a new capstone course or senior seminar. All propose an array of student surveys and student retention studies. Few, however, indicate how assessment is to be locally coordinated or supported. More important, few indicate how the results obtained will be used locally to make improvements.

Because of its legislative origins, negative sanctions, and bureaucratic implementation, assessment in Colorado appears largely compliance-driven. It remains to be seen whether a majority of institutions can convert programs initiated under such circumstances into viable mechanisms for internal revitalization.

Assessment in Colorado appears largely compliance-driven.

Assessment Update
Winter 1989
Volume 1, Number 4

PETER T. EWELL

Assessment in South Dakota

In Minnesota, the state college system has undertaken a "quality indicators" project intended to assess student learning and other dimensions of institutional performance; institutions will propose selected indicators consistent with their missions, and will regularly report on them as part of a quality assurance process.

In Maryland, assessment is included in new accountability legislation accompanying a reorganization of the state's systems of higher education; institutions will be required to undertake "multi-year studies which shall include quantifiable indices of student academic performance."

In Ohio, new legislation has created a Commission on Education Improvement with legislative and board representation; the commission will set goals and assess multiple dimensions of institutional performance, including excellence in undergraduate education.

The program has evolved into a campus-centered effort, intended primarily for local instructional improvement.

Assessment in South Dakota: A State Profile

The story of assessment in South Dakota illustrates many of the dynamics of any state initiative. Begun as an accountability mechanism, the program has evolved into a campus-centered effort, intended primarily for local instructional improvement.

In 1984, the South Dakota Board of Regents adopted Resolution 32-1984 requiring all six public four-year institutions to engage in a program modeled on Northeast Missouri State University's value-added approach. General education was to be assessed through a test-retest process for entering freshmen and second-semester sophomores, using the ACT Assessment and the Objective Form of the ACT-COMP. In addition, all graduating seniors were required to take nationally normed examinations in their major fields. Test scores would not be used to bar students from graduation nor would

Assessment Update
Summer 1990
Volume 2, Number 2

BEST COPY AVAILABLE

they be made a part of their permanent records. To support the effort, the resolution stipulated, students would be charged an additional fee.

Not surprisingly, statewide testing was not warmly received at the campus level. Although intrigued in the abstract by the emerging concept of assessment, administrators and faculty alike were apprehensive about a board-mandated program. Fees to pay for tests unrelated to their programs rankled many students and exacerbated already serious motivation problems. As the board sought to overcome growing resistance, a number of compromises emerged. First, the initial testing program was cut back—particularly in its major field areas: testing in the spring of 1985 was limited to the NTE exam in education, the EIT exam in engineering, and selected GRE major field examinations. More important, to meet fears about public disclosure and inappropriate comparison, the board agreed to a three-year moratorium on releasing obtained scores.

Not surprisingly, statewide testing was not warmly received at the campus level.

Perhaps the most significant move, however, was formation of a statewide testing committee with representatives drawn from each campus. The committee's role was crucial in the program's first three years—both in solving substantial day-to-day problems of implementation and in gradually evolving a more comprehensive, improvement-based concept of assessment as the program's prime intent.

By mid-1987, it was clear that the notion of statewide testing needed a second look. Inflexible test dates, high costs, and inadequate staff for data analysis all severely limited the utility of the data obtained, and made for nearly unworkable administration. In its June 1987 report to the board, the committee recommended major changes to simplify the program and to allow participating campuses greater flexibility.

Ownership of assessment moved from the regents to the campuses and the content of the mandate shifted from requiring individual testing of all students to requiring that instructional programs be systematically evaluated by their own institutions. The newly developed program, launched in 1987-88, continued to require assessment of common areas such as general education, the major field, student satisfaction, and basic skills remediation. But consistent with emerging practice in other states, it allowed for considerable institutional discretion in selecting instruments and approaches.

For South Dakota this policy turnaround has produced the pattern of uneven development typical of most campus-centered state approaches. Some institutions have moved ahead quickly to develop vital, innovative approaches based on locally developed (often qualitative) instruments and techniques; here the program has shown important curricular and policy benefits. Others continue to administer and report on the standardized tests originally mandated. A parallel result has been an evolution of state purposes in assessment toward a more integrated, comprehensive evaluation effort embracing program review, enrollment analysis, and support for accreditation. Both, most observers feel, have been positive developments.

Assessment Update
Summer 1990
Volume 2, Number 2

Survey of State Assessment Initiatives

Few states at this point have nothing to say about assessment.

Developments in state policy occur so quickly that, despite an active "grapevine," it's often hard to get a comprehensive picture of what is happening. Therefore, I'm particularly pleased to have been involved in a new survey of state assessment initiatives, undertaken as a joint venture by the Education Commission of the States (ECS), the State Higher Education Executive Officers (SHEEO), and the AAHE Assessment Forum.

We administered the survey to academic officers of state governing or coordinating boards and obtained responses from all but four of the 52 agencies (including Puerto Rico and the District of Columbia). In some cases, state board respondents were not fully informed about system-level practices, but most provided excellent summaries of what was happening. Many attached copies of current assessment legislation, policy statements, or guidelines. Overall, from both the volume and the quality of the responses, it was clear that much has occurred since ECS undertook a similar survey about two years ago. Space precludes a full reporting of results, but here are some of the highlights:

- Few states at this point have nothing to say about assessment, and 27 have identifiable "assessment initiatives" in the form of legislation or board policies. An additional six or seven say that they "encourage" or "provide active leadership" by sponsoring conferences, providing grant support, and conducting similar activities. Only nine states report nothing in place and nothing planned.
- State interest in assessment shows no signs of "topping out," and 28 of 43 respondents believe that the salience of assessment will increase in their states, while 15 expect interest to "remain about the same"; 10 of these 15 already have active policies in place.
- Most new state initiatives are board policies that allow institutions considerable flexibility. Of the 27 reported initiatives, 18 follow the policy pattern pioneered by such states as Virginia and Colorado. Each institution is required to develop a local "assessment plan" consistent

Assessment Update
Spring 1990
Volume 2, Number 1

But state leaders should not be too quick to dismiss South Dakota's adventure with testing as without policy benefit. Board and campus representatives agree that their experience with statewide testing, though ultimately negative, helped mobilize initial interest in assessment, provided an occasion for serious ongoing discussion, and helped crystallize conclusions about what a good program might look like.

with mission and to report results periodically. None of the newer initiatives was established by explicit, targeted legislation (such as Colorado's HB1187).

- New legislation, where present, includes assessment as part of a more comprehensive "reform" package. Typical of some half-dozen reform bills passed last year are "cutting edge" legislation in South Carolina and SB140 in Ohio. Both call for institutional assessment as part of more extensive institutional reporting and targeted investment. In other states, comprehensive board initiatives fit assessment policy into a wider reform or study package (for example, statewide strategic plans in Connecticut and Idaho and reviews of undergraduate education in Illinois).
- Common standardized assessment measures remain rare. Common cognitive outcomes assessments for college students are in place in only four states (Florida, Georgia, New Jersey, and Tennessee), all but one of which have had them for some time. They are under consideration in only one more (Wisconsin). Three states explicitly reported having recently considered (or tried) common testing and having rejected it. Common basic skills testing is more frequent. Nevertheless, only four states reported using a common instrument, and three more require institutions to choose among a designated set. Four states use periodic common surveys, and 12 reported comparative statistics and student retention.
- States asking for assessment do not generally pay for it. Only six of 27 formal initiatives have a substantial funding base that provides institutions with new dollars to support assessment. The majority of these are programs that have been in place for some time (in Florida, New Jersey, Tennessee, Texas, Virginia, and Washington). An additional four or five report total statewide investments ranging from $25,000 to $500,000 for pilot studies or exploratory activities. Another half-dozen have funded selected institutions through state incentive or categorical grant programs.
- The public posture remains "improvement," not "accountability." Of states responding to a question about assessment's "primary purpose," 31 of 39 reported its purpose to be institutional or curricular improvement, while six considered assessment's prime purpose to be establishing more uniform academic standards. Only two reported that the main purpose is to "demonstrate the credibility of higher education to the legislature and the public," but fifteen reported "accountability" is one of several "major" goals of their programs.

These results provide little evidence of a diminishing state role in assessment, but they do suggest a more deliberate, reasoned approach, with less pressure for "quick fixes" and common comparative solutions. As an academic vice-president with whom I recently talked in Virginia put it, "The good news is that assessment isn't going to go away; the bad news is that assessment isn't going to go away."

"The good news is that assessment isn't going to go away; the bad news is that assessment isn't going to go away."

Mandated Assessment: A Look in the Cultural Mirror

A common consequence of travel is to see familiar conditions reflected in alien circumstances.

A common consequence of travel is to see familiar conditions reflected in alien circumstances. As Trudy Banta also reports (see *Editor's Notes*), we had this opportunity last month, in the form of closely spaced visits to Europe and China. In both places, although in very different ways, an institutional assessment movement is unfolding: and in both places, as in the United States, the process of institutional development is being conditioned by new and controversial actions by government authorities. Some of these parallels in China form the substance of this column.

In China, "evaluation" is primarily a process of course-level assessment. Beginning in the mid-1980s, a few Chinese universities began employing end-of-course evaluation forms somewhat similar to those in common use at many U.S. institutions. Dominating these questionnaires were familiar forced-choice items on instructors' performance, course materials, and classroom environment. By 1987, when I visited China as part of an initial Sino-American Symposium on Evaluation in Higher Education, several Chinese universities were experimenting with different approaches and were looking for new ideas. While familiar tensions existed between faculty and administrators, because of uncertain motives and murky proposals for using the information, the central government was clearly not a player. Moreover, because evaluation results could not be used in real personnel decisions (Chinese faculty are essentially "tenured" at their institutions, on the basis of their initial assignments), evaluation's objective could be more convincingly advanced as the improvement of teaching.

Our latest visit showed considerable evolution of this original, voluntary process. At Peking University, one of the pioneers of the Chinese evaluation movement, traditional questionnaires are increasingly supplemented by newer items on student behavior (such as amount of out-of-class work done as part of the course) and by ratings of classroom teaching, prepared by peer faculty. Quantitative summaries are provided to deans and department chairs for comment and action. We were a bit dismayed by a tendency to overquantify (Chinese practitioners habitually reduce an array of questionnaire

Assessment Update
Winter 1990
Volume 2, Number 4

BEST COPY AVAILABLE

items to a single bottom-line score, through a process of summation and item weighting), but we were impressed with the substantial progress that had been made.

A new wrinkle in evaluation was the presence of the central government, through its State Education Commission (SEC). Responsible for China's 97 "key" universities (those that receive funding directly from the national government, rather than from municipal or provincial authorities), the SEC functions somewhat like a governing board in setting policies and establishing national standards and procedures. This fall, it initiated a pilot program of reporting on institutional evaluation, with plans for making the process mandatory for all institutions in the near future. Mr. Wang Jisheng, deputy director for higher education of the SEC, outlined its main provisions. In concept, the initiative appears broadly consistent with the majority of our state assessment mandates, requiring institutions to engage in a defined process with visible results but leaving the details of what and how to measure up to local authorities. Indeed, the process was quite consciously modeled on U.S. accreditation practice.

Cross-cultural confirmation in this respect is both reassuring and sobering.

In discussing his initiative, Mr. Wang raised some familiar dilemmas of state action. Foremost among them was the need to get all institutions moving, without at the same time stifling appropriate diversity and local ownership. This is a particularly relevant problem in China, where most institutions are highly specialized, often serving only a single industry, profession, or discipline cluster. Embedded in this dilemma is the common problem of uneven development: already, at the conference we attended, there were fears on the part of some of the more advanced institutions that they had progressed beyond the SEC mandate and would need to restructure their current programs to fit the new rules. Also apparent were familiar difficulties regarding the specificity of evaluation guidelines. Institutions wanted levels of generality that allowed considerable latitude, but at the same time they were pressing for clearer signals about what the SEC really wanted. The very openness of the process appeared threatening because its intent was so unclear.

Lurking beneath all this as a complicating factor is the Chinese penchant for ranking. Evaluation practice at the local level rarely emphasizes the use of discrete data items as independent indicators of strength and weakness. Instead, they are used to sort courses into summative categories: "excellent," "adequate," or "unsatisfactory." If such rankings are contemplated for institutions, there are clear disincentives to objective reporting. At the same time, existing institutional prestige rankings are firmly ingrained. Given these predispositions, as well as a historically authoritarian pattern of governmental-institutional relationships, it is hard to imagine that a more uniform evaluation system will not be established eventually.

Whatever the eventual outcome, it was interesting to see—in a new form, halfway around the world—some of the same assessment dilemmas that our states are currently facing. Cross-cultural confirmation in this respect is both reassuring and sobering.

Assessment Update
Winter 1990
Volume 2, Number 4

A New Approach to Quality?

Minnesota Q-7 emphasizes as hallmarks of quality a set of explicit, often surprisingly traditional instructional processes.

For approximately a decade, "total quality management" has been a hot topic in American business and industry circles. Based on the management concepts of W. Edwards Deming and Philip Crosby, among others, TQM, as it is called by its proponents, is based on the intertwined concepts of ongoing incremental improvement, top-level leadership's commitment to distinctive organizational goals, and the establishment of continuous worker-controlled feedback mechanisms for monitoring performance at all levels of the organization. Given the fact that businessmen sit prominently on both higher education governing boards and in many state legislatures, it was only a matter of time before TQM emerged as a force in higher education assessment and accountability. Early last fall, the Minnesota State University System fulfilled this latent promise with the report of its blue-ribbon commission on access and quality, Q-7: *Quality on the Line*.

Called into existence in January 1990, the seventeen-member commission is drawn chiefly from Minnesota's business and professional community. Facing the group was a familiar charge: addressing quality while maintaining the system's historic access role in an increasingly constrained resource situation. Over the past decade, investment has failed to keep pace with a 50% expansion in enrollment. At the same time, an expanding array of knowledge and skills has appeared necessary for the state's workforce to remain economically competitive. Certainly, these conditions are not unique to Minnesota, and the creation of a blue-ribbon group to address them has been the response in many states. Like such highly visible recent bodies as the Select committee on Higher Education in Texas, Arizona's Task Force on Efficiency, Effectiveness, and Competitiveness, or Virginia's Commission on the University of the 21st Century, the Minnesota commission collected background data, conducted extensive interviews, held hearings, and visited campuses. Like these other bodies, it evolved an accountability approach whose centerpiece is a set or explicit indicators of performance. Finally, like many others, the Minnesota commission noted a familiar litany of hoped-for baccalaureate outcomes, including readiness for college, critical thinking and

Assessment Update
January-February 1991
Volume 3, Number 1

BEST COPY AVAILABLE

problem solving, global awareness, multicultural perspective, scientific literacy, readiness for work, and ethical behavior.

But whereas parallel efforts in other systems stressed issues of governance, standards, and resource utilization in their proposed quality-indicators systems, Q-7 emphasizes as hallmarks of quality a set of explicit, often surprisingly traditional instructional processes. For higher-order thinking, for example, mastery is indicated by completion of a senior capstone project or thesis, defended before a jury of faculty members or practicing professionals. Most other indicators have a dual focus, appearing to require both a direct assessment of outcomes and a demonstration that students have been exposed, through regular coursework, to a given body of knowledge or experience. The proposed indicator for global understanding, for example, requires both foreign language study or study abroad and "the ability to articulate the interrelationships of world economics, environment, geography, history, politics, religion, and the arts."

As in many other state- and system-level assessment or quality improvement initiatives, the Minnesota system's seven constituent institutions will be given considerable latitude in developing their own responses. In the coming year, each institution will propose in detail how it plans to collect information and inform improvement consistent with the Q-7 indicators. But in its most intriguing and innovative section, the commission's report clearly signals, in five quality principles, the philosophy that ought to govern each institution's response. Taken directly from the tradition of TQM in industry, these principles include top leadership's commitment and visible resource commitment to the process of quality improvement; inclusive participation by all the system's employees; sensitivity to the needs and demands of customers, such as students, parents, and employers; a working definition of quality, not as an endpoint to be achieved but rather as a process of continuous improvement; and a heavy emphasis on staff training and development.

Customer focus . . . will require far more information about employers' needs, students' goals, and goal fulfillment than traditionally has been the case.

At first glance, these principles seem nothing more than a reworded version of some familiar tenets of good practice in higher education planning. But a second look, I believe, yields an interpretation far more radical with respect to both institutional response and information requirements. Customer focus, for example, strongly and healthily challenges assessment's traditional assumption that faculty are the sole arbiters of academic quality. Meeting this injunction will require far more information about employers' needs, students' goals, and goal fulfillment than traditionally has been the case. Employees' involvement—although it clearly embraces faculty participation at the unit level—also suggests that more than a purely academic dialogue about quality is intended. Together with top-level leadership's commitment, it suggests aggressive institutional leadership to foster active teamwork and create a campus culture of collective responsibility for students' success. Finally, increased investment in staff training implies a faculty development effect that is likely to go well beyond current efforts in inclusiveness and emphasis. Consistent with TQM, it suggests stressing such topics as classroom research for continuous self-monitoring and faculty teamwork to improve the instructional links among sequential or related courses.

Intended or not, these principles suggest considerably more than "reform as usual" in higher education.

Altogether, Q-7 presents a significant challenge to an underfunded, historically conservative university system. If its rhetoric of quality is eventually mirrored in campus-level action, the result may be a transformation in the way we think about assessment and accountability. The effort is certainly thought-provoking and well worth watching.

Q-7 presents a significant challenge to an underfunded, historically conservative university system.

Assessment Update
January-February 1991
Volume 3, Number 1

PETER T. EWELL

Assessment in Hard Times: A Tale of Two States

In 1986 New Jersey and Virginia emerged as front-runners in the new game of state-based assessment. Both mounted generously funded, well planned efforts, and in the coming years both would achieve substantial recognition for their alternative policy models (see this column in the Spring and Summer 1989 issues). Now, both states are in deep financial trouble. And how assessment has fared in each is worth a second look.

New Jersey's approach took the high road. Centered on the controversial and innovative General Intellectual Skills (GIS) examination, the state's College Outcomes Evaluation Program (COEP) remains unchallenged as the most comprehensive state-level assessment initiative attempted to date. From the outset, the program encountered substantial campus opposition—principally centered on the GIS exam. Indeed, for most participants, the GIS was COEP, despite the program's far more extensive (and campus-centered) components embracing student learning in general education and the major field, student development, research and creative activity, and campus contribution to the community.

Also unsettling for many campuses was the perception that COEP was part of an overall strategy for change on the part of the state's Department of Higher Education. While this approach promised campus leaders greater resources and the flexibility to manage them effectively, it also demanded greater accountability; and much of the available new funding was being invested in state-run categorical programs rather than being built into each institution's base. By 1989 increasing budget pressures made this strategy problematic. Resulting staff cutbacks at the department also rendered the original COEP designs infeasible: remaining staff were unable to react to an increasing volume of campus reports or to further encourage local assessment development. The cornerstone of the program was seen to be the GIS exam, and all remaining resources were concentrated on putting it into place.

New Jersey's approach . . . remains unchallenged as the most comprehensive state-level assessment initiative attempted to date.

Assessment Update
January-February 1992
Volume 4, Number 1

BEST COPY AVAILABLE

The eventual implementation of GIS in the spring of 1990 (after three years of development) was a technical triumph but a mixed success politically. Resources to support its first administration were denied by the legislature but were obtained from the governor's contingency fund. Several campuses boycotted the program and others complied under protests. Despite these difficulties, over 4,300 examinations were administered and scored and the results publicly reported. Much, however, had depended on strong departmental leadership and solid gubernatorial support. With a change of both in 1991, the program's future became doubtful. Again the legislature refused to fund the program, and this time the decision was made to suspend it.

While GIS was administered as planned in the spring of 1991, results were withheld—partly for budgetary reasons but probably also because their content was politically unpalatable. COEP is now officially "on hold"; some department officials want to revive it if times get better, but the program's rebirth in its original form is improbable. Meanwhile, campus-based assessment in New Jersey never really got started, and in the absence of a strong locally centered reporting mandate, many incipient programs themselves fell victim to budget cuts in 1989-1991.

Several campuses boycotted the program and others complied under protests.

Virginia, meanwhile, had taken a different road. In 1987, prompted by the legislature, the State Council on Higher Education (SCHEV) required each campus to develop its own assessment plan. While broad guidelines were provided, institutions were encouraged to adopt assessment approaches that best fit their missions, curricula, and student clienteles. On approval by the council, institutions were granted additional resources to implement their plans—resources that were later built into each institution's base allocation. Each biennium thereafter, institutions were required to report results comprehensively, addressing methods used, what was found, and what actions were taken as a result.

SCHEV's plan was also controversial, despite its campus-centered approach. Institutions complained about insufficient resources and short timelines for response, and in a state where institutional autonomy is the prevailing culture, they were also profoundly suspicious of state motives. But by the end of 1987, each had a plan in place and resources dedicated to carrying it out. Subsequent biennial reports to SCHEV in 1989 and 1991 revealed uneven response but for the most part a slow and steady progress in building a set of credible local processes.

By the end of 1989, Virginia was also facing severe budgetary problems, affecting state assessment policy in at least two ways. First, the decision to include assessment funds in each institution's base meant that institutional leaders were tempted to reallocate these funds to higher local priorities. Fear of this kind of response was precisely why state officials in New Jersey had decided on a far more directive approach in the first place. At the same time, the immense diversity of campus assessment activities made it hard for SCHEV to communicate exactly what assessment dollars were buying. SCHEV dealt with the first condition by privately, but severely, warning campus leaders off this course of action-recalling their ultimate authority

Assessment Update
January-February 1992
Volume 4, Number 1

to deny campuses access to discretionary funding in the absence of an "approved" assessment program. The second is a continuing problem.

Comparing the two states' approaches, Virginia's has certainly proven the more robust. While on many (and perhaps even a majority of) Virginia campuses assessment is still struggling, state action in many cases has proven an effective stimulus for gradually developing a valued and valuable local program.

It is useful to speculate what might have occurred had New Jersey's COEP enjoyed two or three additional years of generous funding and political goodwill. With the GIS exam in place, state priorities might have given more attention to local assessment concerns and to linking GIS results with statewide improvement efforts. As it happened, linkages to action and reform came too late to show the real utility of assessment, both to legislators and to campus leaders. But equally tempting when hard times hit might have been a proposal to use GIS results for "performance funding," a concept now enjoying something of a renaissance among state officials. In politics, some have said, timing is everything. It's an observation that seems particularly germane these days for those involved in state assessment policy.

In Virginia state action has proven an effective stimulus for gradually developing a valued and valuable local program.

Assessment Update
January-February 1992
Volume 4, Number 1

Regional Accreditation: The Uncertain Alternative

With two-thirds of the states already in the game, . . . accreditation is clearly playing "catch-up."

From its beginnings, the "public" part of assessment has come at institutions from two directions: the mandates of state government and the requirements of regional accrediting associations. Of the two, the latter has always seemed preferable. Indeed, in the early days, many in the academy believed that if voluntary accreditation could demonstrate a credible track record on assessment, state governments might eventually back off. Recent events in Washington and elsewhere, however, are rendering this option problematic and raise fundamental questions about accreditation's future role in assessment.

Compared to the states, regional accrediting bodies approached assessment gingerly. Five years ago only one, the Southern Association of Colleges and Schools (SACS), specifically referenced assessment in its requirements; by then, over a dozen states had enacted assessment mandates. Now all six regionals require assessment as an identifiable part of institutional reaffirmation. More important, all six expect from institutions approximately the same things: clear statements of intended educational results, explicit evidence regarding the attainment of these results, and a visible institutional mechanism (and eventually a "track record" as well) for using the resulting information. Like the majority of state mandates, accreditation requires no specifically identified assessment methods. Institutions are both allowed and encouraged to do what suits them best.

For most regionals, assessment is a very new activity, and its early dynamics are familiar; institutions with which I've worked under the new North Central Association guidelines, for instance, tend to exhibit the same initial "what-do-they-really-want-us-to-do?" bewilderment as was apparent in the first year of state-based assessment in places like Virginia, New Jersey, or Colorado. With two-thirds of the states already in the game, however, accreditation is clearly playing "catch-up." At the same time, its institution-centered, goal-driven approach is hard to distinguish from the majority of state-based programs—programs that generally involve far higher stakes for the institutions affected. As a result, the impact of new accreditation requirements has

Assessment Update
May-June 1992
Volume 4, Number 3

BEST COPY AVAILABLE

largely been felt by private institutions and in those states that have not made an explicit move in assessment.

Also like the early days of most state mandates, the real issue is less the adequacy of specific assessment measures than the public credibility of the resulting process. And here the public position of the regionals has eroded considerably. Up to now, a good deal of accreditation's clout has rested upon its role in determining eligibility for the receipt of federal funds. For assessment, the federal connection became important about four years ago, when the Department of Education first required accreditors to explicitly include outcome measures in their institutional review procedures as a condition of recognition. At that time, it appeared that government authorities preferred working through existing voluntary processes to designing and imposing their own. But recent attacks on the Middle States Association by the Secretary of Education and specific proposals by both houses of Congress to write accreditation out of federal eligibility decisions during Higher Education Act reauthorization debates this year seem to signal a new public attitude: accreditation's value and credibility as an accountability device now appears questionable to many in government—with or without outcomes measures.

Accreditation must creatively rethink its role in assessment.

If deemed inadequate to assure accountability, what should accreditation do with assessment? Noting both accountability and improvement in their rhetoric, accreditors have up to now wanted to have it both ways. But accreditation's summative claims have never implied a guarantee of specific educational products. Instead, they have centered on more holistic concerns with the appropriateness of institutional purpose and the adequacy, deployment, and impact of resources consistent with such purposes. By refocusing attention more aggressively on their historic role of "validating" institutional "quality assurance" processes through self-study, accreditors might capitalize on an historic strength and still clearly distinguish what they do in assessment from the actions of governmental authorities. Like the Malcolm Baldrige Award for "good practice" in quality assurance in business and industry, a voluntary process could provide a focus for systematic, continuous improvement of institutional assessment practices. At the same time, it might powerfully complement a growing state and federal role in directly examining the adequacy of higher education's products.

Whatever the mechanism chosen, accreditation must creatively rethink its role in assessment. In an uncertain future, we need it badly. Its core activities of self-study and peer review have a strong tradition, and are visibly owned by the academy; with these two virtues, voluntary accreditation is still among our best alternatives for getting the "public" part of assessment right.

Assessment Update
May-June 1992
Volume 4, Number 3

Developing Performance Indicators for Community Colleges: Evidence from Two States

Distinctive approaches to assessment have evolved for the two-year college sectors in many states.

In the realm of accountability, community colleges have always held a position about midway between K-12 and the balance of higher education. A first reason lies in governance. A majority of community colleges, like schools, are administered by local district boards. On a state level, particularly in the vocational/technical area, the roots of community college accountability and reporting lie in the agencies responsible for secondary education. Not least in the two-year college world, the federal government is prominently present through its direct dollar support for vocational and technical instruction. A second reason for special treatment is that, by their very nature, the products of community college instruction appear more measurable than those of four-year college programs. The primary assignments of job training and basic skills instruction provide little of the "ineffable" to hide behind when it comes to giving evidence.

For these reasons, distinctive approaches to assessment have evolved for the two-year college sectors in many states. These approaches are often more quantitative and prescriptive than assessment methods for four-year colleges are. And in light of more general trends, at both state and federal levels, toward the development of performance indicators, the experiences of two states—Texas and California—are worth a second look.

The linkage of assessment and accountability began in Texas in 1987 with passage of HB 2182, which mandates basic skills testing and evaluation of the effectiveness of remediation programs through required follow-up. The resulting Texas Academic Skills Program (TASP) is based on a common examination designed by National Evaluation Systems and administered to all incoming students statewide as a placement tool. To implement the required follow-up evaluation, statistical indicators were developed to measure subsequent performance in regular college-level work of students who were originally assessed through TASP as deficient.

Assessment Update
July-August 1992
Volume 4, Number 4

In California, accountability came more slowly but was more comprehensive when it arrived. Following several study resolutions on the development of

BEST COPY AVAILABLE

community college performance measures introduced by Assemblyman Tom Hayden, in 1988 the California legislature passed AB 1725. The bill calls for the development of indicators of the educational and fiscal performance of the community colleges, but it left the colleges free to determine the precise content and structure of these indicators.

Both states have relied on a contemporary assumption of assessment policy: properly designed, a single approach might serve both state and institutional purposes. The legislation thus emphasizes the development of technical capacity at the institutional level to generate the required statistics and holds out the promise that large, centralized, state-held records systems might be used to generate information of local utility. But in pursuing these noble goals, some more general lessons about indicators and accountability have emerged.

Technical capacity is critical. In both states, community colleges were ill equipped to respond to their new mandates. Few had wide experience in data collection and analysis, and computing capabilities were often extremely limited. At least as significant was the fact that community colleges had little tradition of collective action, other than lobbying. Common definitions required for meaningful reporting did not exist, and there were few available mechanisms to develop them quickly. State authorities responded with substantial efforts to develop local institutional research capacities. In Texas, federal Carl Perkins funds supported a consortial effort to develop a prototype longitudinal follow-up capability. In California, $375,000 was appropriated to support four pilot projects and a model accountability reporting system. In both states, this early investment in technical capacity has proved critical for an effective response.

Texas and California have relied on a contemporary assumption of assessment policy: . . . a single approach might serve both state and institutional purposes.

But "high-tech" solutions aren't always best. Experience in both states has also suggested that the most complex methods for assembling data are not always the most fruitful. In Texas, after two years of pilot development, the state abandoned direct support of an institution-based student tracking consortium in favor of a more centralized effort that links unit-record enrollment, testing, and high school attendance files. This technically complex approach has not fulfilled its initial promises. In California, pilot campuses have generally found small, flexible databases (often implemented on a microcomputer) more useful in generating the required statistics than solutions of greater elegance that involve investments in large relational databases.

Incentives to institutions need careful balancing. Linking high stakes and measured performance can have negative consequences. If the states are high, campuses will start "managing to the numbers." But if the consequences of not participating are few, measured performance will simply be ignored. Both states are currently struggling with this dilemma. In Texas, proposals are on the table to use statistics about subsequent student success in college-level work to "performance fund" developmental instruction. In California, additional proposals to link community college performance to state allocation have arisen. As indicator approaches become more nationally prominent, this policy dilemma will become more apparent. There is no easy answer.

"Fair" comparisons are tricky. By definition, performance indicators are meant to be compared. But for policy purposes, what constitutes an appropriate comparison? Community colleges in Texas and California differ markedly in size, setting, program mix, and student clientele. These differences have an automatic impact on such statistics as retention and transfer rates, regardless of any actual differences in institutional performance. The experience in both states suggests that comparative performance indicators are best presented in conjunction with a parallel set of contextual statistics that explicitly show how institutions differ. As illustrated by recent debates about federal "student-right-to-know" reporting, this lesson needs a wider hearing.

In some areas a centralized approach may be best.

States really can do some things better. Tensions between centralization and institutional discretion have been present in assessment policy since the beginning. In most cases, institutions have argued successfully that real differences in mission and clientele render common methodologies inappropriate in virtually all domains. Experience in both states, however, suggests that in some areas a centralized approach may be best. In Texas, statewide efforts to link transcript data electronically have proved particularly promising as an alternative to institutional follow-up surveys; similar linkages between graduate files and state employment-reporting databases have been similarly successful. By their very nature, such efforts must be centralized, and they must be based on an efficient, state-administered unit-record reporting system.

This brief look at two-year college accountability in Texas and California suggests, first, that performance indicators are publicly appealing and will arise with increasing frequency as an answer to the accountability issue. It also suggests that state-institution partnerships can work in developing such indicators, given time, funding, and above all a mutual recognition of appropriate roles. In what promises to be a much wider national discussion of performance indicators, we should take these lessons to heart.

Assessment Update
July-August 1992
Volume 4, Number 4

PETER T. EWELL

Institutional Consortia: A Promising Alternative for State Assessment Policy

Regardless of their content, most state assessment approaches up to now have followed a common administrative recipe. A mandate is enacted by legislature, board, or executive fiat, and direct responsibility for implementation promptly turned over to an appropriate arm of state government—generally the state's coordinating or governing board for higher education. Active management of the process becomes a line responsibility for designated coordinating board staff, including the development of appropriate regulations and reporting requirements, sharing information, and assisting in the development of assessment expertise.

Most state assessment approaches up to now have followed a common administrative recipe.

Assessment, under this approach, becomes a regularized bureaucratic function similar to state-level program approval or master planning. Here, the primary organizational relationship is between each institution individually and the state capitol, with concomitant difficulties of politics and communication. But the recent experiences of several states—most notably South Carolina—suggest that it doesn't need to work this way. Voluntary consortia of institutions, encouraged and in some cases actively supported financially by state government, can discharge many of the functions normally fulfilled by government agencies, and at the same time may far more effectively promote the kinds of inter-campus collaboration and problem solving most likely to lead to effective institutional practice.

The South Carolina Higher Education Assessment (SCHEA) Network was formed in 1988 to help colleges and universities across the state develop sound campus-based assessment practices. Initially established through a FIPSE grant and funding from the South Carolina Commission on Higher Education (with matching funds provided by the host institution, Winthrop College), the SCHEA Network currently has over 40 members including both public and private institutions. As expected, primary concerns of the Network are driven by common needs—in this case, the requirement for member institutions to respond to SACS institutional effectiveness criteria and the need for public institutions to respond to parallel, more specific, reporting requirements mandated by the state commission.

Assessment Update
September-October 1991
Volume 3, Number 5

BEST COPY AVAILABLE

Included in its regular activities are:

- An annual conference providing nationally known speakers, an opportunity to showcase the work of Network members, and the chance to share experiences with colleagues
- A campus workshop/consultation service based on expertise within the state
- A semiannual newsletter containing excellent articles on assessment practice (both technical and organizational) and updates on member campus activities
- A number of published resources, including an assessment bibliography (currently, in my opinion, the most comprehensive available anywhere), and a useful *Beginner's Guide* to higher education assessment (most recently used as the basis for several well received workshops at the national AAHE Assessment Forum).

Important communications functions can be carried out among colleagues, with state representatives sitting at the table as peers.

Many of these activities, of course, are of the type directly or indirectly provided by state agencies in mandate states. In New Jersey, for instance, detailed technical manuals covering a range of assessment procedures were produced directly by the Department of Higher Education through its College Outcomes Evaluation Program. Bibliographies and directories of state practice, moreover, have been developed by almost a third of the states currently active in assessment. But other states have partially followed South Carolina's approach in developing voluntary government-institutional partnerships. In Virginia, for example, an annual statewide assessment conference is planned and implemented by an independent group of campus assessment coordinators who are financially supported by the State Council. And in Texas, a consortial approach to developing longitudinal student data systems was supported by Coordinating Board funds, but operates as an independent membership organization.

From both state and institutional perspectives, such an approach has several advantages. From the state's point of view, important communications functions can be carried out among colleagues, with state representatives sitting at the table as peers. In South Carolina, Commission staff report that they are far less directive about assessment than their counterparts in other states because the Network itself serves as convener. At the same time, they feel, technical assistance functions provided by the Network can be more efficiently and effectively run by practitioners themselves. Institutions, in turn, can direct consortial activities more flexibly toward areas that best meet their needs. But they also have the informal opportunity to query Commission staff about proposed approaches and to inform them more fully about ongoing campus activities.

Over the years, many state assessment associations have formed naturally as a means of mutual assistance, reassurance, and protection. South Carolina's experience suggests that they may also constitute an intentional approach to state assessment policy. Certainly it's an approach worth greater scrutiny.

For further information on the SCHEA Network, contact Dr. Reid Johnson, 210 Tillman Hall, Winthrop College, Rock Hill, SC 29733.

Assessment Update
September-October 1991
Volume 3, Number 5

PETER T. EWELL

A New Federal Role in Assessing Outcomes

Up to now, the story of assessment and accountability has been confined largely to the states. With investments in public higher education of anywhere from one-quarter to one-third of general fund revenue, state governments have a direct and compelling interest in how and on what these dollars are spent. At the same time, through direct regulatory and governance mechanisms, they possess the necessary tools for ensuring institutional compliance. Lacking both, the federal government, for the most part, has remained aloof. Now it appears that this situation is changing.

Several recent developments mark this shift of posture. The most visible, perhaps, is Public Law 101-542, finalized by Congress after considerable conference-committee negotiation in January, under the revealing rubric of students' "right to know." As negotiated in conference, the law represents a typical congressional conglomerate of concerns about higher education. Basically, it contains three major provisions. Under Section 103, all postsecondary institutions will be required to track and "make available" rates of graduation or program completion for degree-seeking students. The law calls for institutions to begin tracking graduates by July 1991, with the first statistics to be made public by July 1993. It also calls on the Secretary of Education to report back to Congress by August of this year on the feasibility of reporting graduation rates by field of study, pass rates on licensure and certificate examinations in applicable fields, employment rates in fields of training, and "other outcomes as appropriate." While the provisions of Section 103 have the greatest potential impact on most institutions, Sections 104 and 105 reveal the primary sources of legislative concern. The first requires more detailed reporting of performance and completion rates for student athletes; the second requires regular public disclosure of statistics on campus crime.

None of the concerns embodied in this legislation, of course, are particularly new. Indeed, its reporting requirements strongly resemble earlier "track record disclosure" regulations applied to federally funded two-year occupational/technical programs and their massive predecessor, the Vocational Educational Data System (VEDS). Both, it is important to stress, remained

The federal government has remained mostly aloof from assessment, but now that appears to be changing.

Assessment Update
May-June 1991
Volume 3, Number 3

BEST COPY AVAILABLE

largely unimplemented because of the sheer technical inability of most institutions to generate the necessary statistics. Like track record disclosure, moreover, the new law demonstrates an overriding concern for consumer protection. This, too, is an old theme, dating back to the tone set by former Secretary of Education William Bennett's frequent calls for colleges and universities to justify high costs by demonstrating commensurate benefits to students and their parents. As consumer protection, the new law appears at first to govern a direct relationship between student and institution: there is no apparent need to report statistics to the Department of Education, in standard or comparable form. Given this history, many in higher education may justifiably say that "this too shall pass." Bennett's exhortations, after all, remained only that; VEDS and track record disclosure, in many ways, collapsed of their own weight.

But additional national developments reinforce the perception that something real is happening. Most developments now center on the need to develop national indicators of progress in education, as a result of last year's "education summit," which brought together in Charlottesville, Virginia, President Bush and all fifty of the nation's governors. While only one of the six national goals adopted by the governors addresses higher education (goal five, on literacy and lifelong learning), the target is clearly degree productivity and cognitive outcomes. The last two objectives under goal five state, in short, that there should be "substantial increases" in the proportions of students who "persist in college and complete their programs" and in the proportion of college graduates who can "think critically, communicate effectively, and solve problems."

Federal interest in collegiate outcomes will continue to grow.

How to assess progress on these goals is, of course, an overriding concern. After almost a year of work, the National Education Goals Panel, under the leadership of Governor Romer of Colorado, is preparing to deliver its recommendations for appropriate and attainable national targets within each goal and for monitoring their attainment. Current recommendations to the Romer panel by specially convened groups of policy researchers are unlikely to advocate any national test, but there has been much discussion in those meetings of the need to develop better overall measures of higher-order thinking. Prominent in these discussions, for example, has been the development of an examination that could embrace the kinds of higher-order skills appropriate to at least the first two years of college and that would involve samples large enough to provide credible state-level results. Also clear from these discussions is the need for visible quantitative indicators for tracking overall progress over time and for comparing performance across states in the realm of college-level persistence and degree completion.

Again, it is too early to tell where all of this will come out. It seems likely, in the short run, that the Romer panel will endorse current state-level efforts to assess collegiate learning by requesting each state to report what it is doing. But it seems an equally compelling conclusion that federal interest in collegiate outcomes will continue to grow. If, as the preamble to the new law suggests, outcomes reporting is tied directly to institutions' eligibility for federal financial assistance, the stakes may be high indeed.

Assessment Update
May-June 1991
Volume 3, Number 3

PETER T. EWELL

"Program Excellence" in Ohio: An Indirect Approach to Assessment

Ohio is among a minority of states that has avoided an explicit assessment mandate for higher education. But for nearly ten years, the state attempted to encourage assessment indirectly through an incentive funding program that was one of the first of its kind in the nation. Because proposals to link incentive funding and assessment are becoming more frequent, Ohio's experience with this approach is worth a second look, and a recently completed statewide evaluation of the program provides an excellent opportunity for this reexamination.

Initiated in 1983, the Program Excellence program in Ohio was part of a larger incentive funding effort, Selective Excellence, that ultimately distributed over $200 million to public higher education institutions in the form of addition-to-base competitive or categorical grants. Like similar programs in New Jersey, Virginia, Colorado, and other states, Program Excellence relied on competition among state institutions to stimulate quality and innovation in undergraduate education. One-time awards of up to $150,000 were given to submitted programs after a rigorous external review process that emphasized outcomes data on student placement and performance. In addition to directly funding improvements, the program's architects at the Ohio Board of Regents hoped that this connection would stimulate institutions to initiate or enhance their efforts in student assessment without the need for a mandate. Indeed, by the mid-1980s they were visible proponents of this "indirect approach" as a policy alternative to the mandates that were then becoming prominent in other states.

Like the effects of any decentralized approach, the impact of Program Excellence on campus assessment practice varied widely. First, because the unit of analysis was the individual academic program, it was quite possible for this approach to induce considerable enhancement of particular program-level assessment practices without affecting similar efforts in other programs at the same institution, or the effects of the institution as a whole. Indeed, at many institutions, access to the Program Excellence competition was limited to programs that could already demonstrate excellence through an established

Ohio attempted to encourage assessment indirectly through an incentive funding program that was one of the first of its kind in the nation.

Assessment Update
January-February 1993
Volume 5, Number 1

BEST COPY AVAILABLE

assessment effort, which was most often the result of a need to meet the requirements of special accreditation.

At other institutions, as intended, campus leadership took advantage of the opportunity provided by Program Excellence to link the application and review process to a wider institutional assessment agenda. At Columbus State Community College, for instance, a veterinary technology program funded under Program Excellence was designated the institution's "lead program" on assessment and was required to share the results of its efforts as a model for other departments.

In other cases, the stimulus provided by Program Excellence allowed previously moribund review processes to be rekindled. At Clark State Community College, for example, an administration-initiated program review system had fallen into disuse because nothing was at stake; incentive dollars provided through Selective Excellence gave institutional leaders an opportunity to restart the process on a new basis, with something real to allocate.

> *At present, the future of Program Excellence is uncertain—together with the rest of Selective Excellence.*

Finally, Program Excellence directly supported experiments with new assessment techniques. Among the most innovative was the "beeper study" on student use of time, conducted by Karla Schilling at the Western College Program of Miami University, in which students were equipped with individual electronic paging devices and requested to record their current activities in a diary when "beeped" by the study directors at pre-defined intervals.

From a policy standpoint, however, the impact of Program Excellence on institutional assessment practice in Ohio was problematic. The availability of incentive funds reinforced existing or incipient assessment efforts far more than it stimulated new ones. At the same time, despite the program board's success with a derivative Fund for the Improvement of Postsecondary Education (FIPSE) project intended to pilot a wider voluntary review process, Program Excellence functioned largely as an add-on to existing state policy mechanisms. As a result, institutions could opt out if they did not see a direct benefit. Indeed, several of the state's larger universities ceased competing for Program Excellence because the need to document outcomes rendered the effort "too expensive" in relation to the potential incentive dollars that might be received. Finally, with the exception of the FIPSE project, few statewide mechanisms were available from which to generalize or to share institutional experiences. Unlike Virginia and South Carolina, which initiated vigorous consortial efforts endorsed and supported by state authorities, Program Excellence principally resulted in an isolated set of program-level experiments.

At present, the future of Program Excellence is uncertain—together with the rest of Selective Excellence. The program received no funding for the current biennium, largely as a result of the state's increasing fiscal problems and a consequent need to reevaluate state policy alternatives. Proposals for the future currently on the table in Ohio, as in many states, include strengthened accountability and evaluation mechanisms and a new set of

Assessment Update
January-February 1993
Volume 5, Number 1

incentive funding programs that are more closely tied to clearly identified state needs. As these proposals are weighed, the lessons of Selective Excellence are proving helpful to the formulation of a new generation of policies. Among the most prominent of these lessons are (1) the power of positive incentives over negative sanctions to induce desired institutional behaviors, (2) the need to actively connect the results of assessment to stakes that are important enough to matter to both institutions and programs, and (3) the need to build effective mechanisms for sharing isolated institutional experiences with assessment and improvement on a statewide basis. As a pioneering state policy effort, Program Excellence in Ohio can claim considerable credit for initiating informative assessment conversations. Its successor, in turn, may signal a new generation of state policies in which assessment and funding are more tightly intertwined.

As a pioneering state policy effort, Program Excellence in Ohio can claim considerable credit for initiating informative assessment conversations.

Assessment Update
January-February 1993
Volume 5, Number 1

Performance Indicators: A New Round of Accountability

One major difference in accountability's "new look" is its considerably wider domain.

Many have noted that state higher education policy agendas have now tilted strongly toward accountability, with marked impacts on state approaches to assessment. Another manifestation of this shift in many states, though less openly discussed, is the rapid development of statistical performance indicators as a quick answer to higher education's accountability "problem." In the last three months, under the auspices of an Education Commission of the States project aimed at charting and analyzing these developments, I've had an occasion to look at what's happening in ten states and to make some preliminary generalizations. In my column this issue I'd like to look briefly at the broader context for the development of state performance indicators, to examine some themes of implementation, and to describe specifically what these schemes require.

One place to start is by asking how the current accountability conversation resembles the one that launched assessment some seven years ago. One major difference in accountability's "new look" is its considerably wider domain. While the emergence of state interest in assessment in the mid eighties signalled a concern with "quality" in opposition to more traditional elements, such as access and efficiency, current requirements demand both. Partly this is a matter of changing context. Most states initially approached assessment in good (or at best, fiscally neutral) times, and could afford to consider "quality" as an add-on. Now, states recognize, hard times require doing more with less: gaining "quality" without sacrificing efficiency and access. As a result, the current crop of indicators proposals mixes new and old elements. Most appear dominated by a concern with traditionally defined efficiency, but many—like South Carolina's new requirement to report such things as the number of undergraduate students actively participating in sponsored-research activities—reflect a new concern with explicit educational processes and delivery as well.

Assessment Update
May-June 1993
Volume 5, Number 3

One way that the move toward indicators looks like early state efforts in assessment, though, is in the incredible speed of its evolution. Kentucky's SB109, South Carolina's Act 225, and New Mexico's report card bill" look alike because their architects directly and rapidly informed one another in a pattern that might truly be labeled "legislation by fax." Indicators proposals in other states arose in very different ways (Wisconsin's, for instance, was an indirect product of a "blue ribbon" task force on faculty compensation!), but because of both a small body of available state-level data and limited technical possibilities for manipulating it, these proposals have ended up with a good deal in common. Most emerging indicators systems, for example, contain in total some 15 to 20 distinct data items collected by the state higher education governing or coordinating body and are reported in the form of direct comparisons among institutions or sectors. In contrast to state assessment initiatives, moreover, the majority of new indicators schemes have a visible connection to funding. Though only a few states (Texas and Arkansas, for instance) appear to be joining Tennessee in an explicitly performance-based approach, many (like South Carolina or New York) do link reporting compliance with base funds eligibility or access to a variety of special purpose funds. Also in contrast to most early assessment initiatives, state authorities have been less willing to change proposed requirements when faced with initial (and often vociferous) protests from institutions. The predominant attitude has been hard line: that higher education has up to now "escaped" the kinds of real accountability measures applied to other state agencies and that this situation is no longer tolerable.

Higher education has up to now "escaped" the kinds of real accountability measures applied to other state agencies and that this situation is no longer tolerable.

What, finally, are these states collecting indicators about? Probably the most common single item is graduation/retention rates, often compiled via existing state-level unit record systems. A second frequently included component is graduate placement, either in field-related employment or in further postsecondary education. Somewhat similar are requirements to provide "linkage" data across educational sectors such as community college to four-year transfer rates, or reports directed toward primary feeder high schools. Reflecting concerns with instructional quality, a third category of indicators directly examines instructional practices; several states, for instance, require institutions to report the proportion of lower-division credits delivered by full-time faculty. Finally, some indicators lists do indeed contain cognitive outcomes, though most states have continued to avoid proposals for expensive common testing. Wisconsin's, for example, contains a proposal for end-of-sophomore-year testing using the ACT CAAP, while a number of SREB states are engaged in exploratory conversations with ETS about administering a version of New Jersey's task-based General Intellectual Skills test discontinued by that state some two years ago.

These preliminary findings signal a return of legislative interests in simple "bottom-line" answers about higher education's performance. With times as tough as they are, these questions will be much harder for us to duck than in their first "assessment-mandate" incarnation. Then as now, we'd best be ready with workable alternatives.

Assessment Update
May-June 1993
Volume 5, Number 3

More Federal Follies

In D.C., ironically, things have never appeared more unsettled.

When *Assessment Update* first appeared some four years ago, we on the editorial board saw little need for a "From the Feds" counterpart to this column. Now, it seems, I spend about half my time describing events in Washington. Not only is this indicative of trends in accountability generally, but it also says something about what's happening at the state level. Particularly with respect to such expensive enterprises as new cognitive assessments for college students, for instance, some states appear visibly delighted to see the federal government doing the heavy lifting. Others are taking a "wait-and-see" attitude-counting on the passage of time to sort out the accountability picture.

In D.C., ironically, things have never appeared more unsettled. Consider, for instance, recent events surrounding graduation-rate reporting. Four distinct efforts on this front are now underway at the national level, each with its own particular "take" on the problem and each proceeding largely as if the others did not exist. First, as most readers will now be aware, final rule making for Student Right-to-Know (SRK), which has been federal law for almost three years now, has been delayed until at least the fall and will occur after an almost unprecedented second national comment period this summer. This means that the first required report under this legislation, due July 1, will be occurring without official rules. Meanwhile, in January the National Center for Education Statistics (NCES) will pilot a format for national reporting of persistence and graduation rates at the institutional level through its annual IPEDS data-gathering process, an effort whose scope and required level of detail go well beyond the original mandate of Student Right-to-Know. In parallel, under the banner of Goal 5.4, a technical advisory panel of the National Education Goals Panel has come out with a first draft of its own format for state-level reporting of graduation and persistence rates that includes part-time as well as full-time students. Finally, as part of the reauthorization process, Department of Education staff responsible for implementing Part H of Title IV of the Higher Education Act have been charged with developing a range of indicator statistics to be applied to institutions that "trigger" investigative attention because of high

Assessment Update
July-August 1993
Volume 5, Number 4

BEST COPY AVAILABLE

default rates on federally guaranteed student loans. This plethora of activity has left most institutions understandably confused—and perhaps privately pleased—by the resulting delays. But behind these developments, I'm afraid, are some long-range implications for national accountability.

The Part H discussion is perhaps the most ominous of these, as it's the first time, to my knowledge, that the word *standards* is being used in a federal context about a postsecondary performance statistic. According to current thinking, institutions tripping the default "trigger" will likely be explicitly tracked on such items as graduation rates. If they fall below a specified standard (state or national has not been decided as of this writing), they will lose their eligibility to participate in federal aid programs. The department already has the authority to use default rates alone to make such a judgment, so why invest time developing additional standards for postsecondary performance? The answer may be that once established, these standards will prove more generally useful. The long gestation period of SRK also fits this interpretation. A recent conversation with a Department of Education official contained the intriguing comment that one reason for the delay is to ensure maximum consideration of the technical properties of a statistic that the secretary appears very interested in using for multiple purposes in future policy, some of which may involve high stakes for institutions.

Part H of Title IV of the Higher Education Act is perhaps the first time that the word standards *is being used in a federal context about a postsecondary performance statistic.*

While the postsecondary graduation rate discussion is well ahead of discussions about a national test, readers may be aware that the latter conversation also has not stood still. Earlier in 1993, NCES brought out its expected request for proposals (RFP) on the development of a national higher education assessment, the result of a series of design conferences on this topic held over the past two years under the auspices of National Goal 5.5 that calls for increases in the ability of college graduates to "think critically, communicate effectively, and solve problems." The RFP calls for proposing agencies to administer a two-year-long process to achieve consensus on the precise content of these abilities and review current and potential measurement technologies that might be used in a national collegiate assessment. At the same time, a third successive technical task force convened by the National Education Goals Panel (the prior two not having come forward with sufficiently forceful recommendations on this topic) recommends development of a "constellation" of national indicators of collegiate achievement (including relevant occupational skills) and is holding hearings on its proposals throughout the spring. Neither of these efforts at this point has the immediacy of the graduation-rate discussion, which requires statistics to be compiled for individual institutions. Both national assessment conversations are confined to development of a "national indicator," without explicit standards, that will be disaggregated, at most, to the state level for purposes of public reporting. But anyone familiar with the history of the National Assessment of Educational Progress (NAEP) will remember that this enterprise began this way as well.

To readers who fear that *From the States* has permanently acquired the sensationalist tone of a gossip column, I apologize. It's merely the way things look at present to one interested observer. But I'd remind any who think that the change of administration means a pause in the emerging federal assessment agenda of one useful bit of history: one of the original authors of the National Governors' Association's "Time for Results" report in 1986 was none other than the then governor of Arkansas.

Assessment Update
July-August 1993
Volume 5, Number 4

Performance Funding in Texas

Guest Columnist: Gerald H. Gaither, director of institution research at the Prairie View campus of the Texas A&M University System.

From the States for this issue is written by my friend Jerry Gaither, who provides a good overview of Texas's "near miss" with performance funding. Other states currently examining performance funding options are Missouri, Minnesota, and Colorado.

— Peter T. Ewell

Texas is now struggling with the task of developing a coherent performance funding system for its public institutions.

Due to many factors, Texas is now struggling with the task of developing a coherent performance funding system for its public institutions. Following the oil bust of the mid 1980s, state spending for higher education failed to keep pace with needs. Over the past eight years, for example, inflation-adjusted appropriations for higher education increased only 1%. Growing external pressure for greater public accountability also encouraged state policymakers to seek sharper focus in public agency performance and to develop specific funding methods to meet these challenges.

In 1987, the "Texas Charter for Public Higher Education" stated that "Incentive Funding shall be established to reward institutions achieving specific goals." The charter further required the Texas Higher Education Coordinating Board (THEC) to develop a multifactor system to distribute such funds, and to recognize various campus efforts through an incentive funding program. Two principles set forth in the charter are *quality* and *access*, and these principles provided the guiding framework for developing the proposed performance-based funding model.

Ethnic and racial issues, clouded in legal difficulties, also had an impact. Several court cases in the 1980s led to redistribution of funds that favored minority institutions. In January 1992, a state district judge ruled that the Texas higher education system discriminated against residents of primarily Hispanic south Texas, a decision which was appealed by the state. Against

BEST COPY AVAILABLE

the backdrop of these principles and events, the state's performance funding plan was developed.

For six years, the performance funding concept lay dormant in Texas, and no funding was provided for its implementation. Two events were primarily responsible for renewing the issue. First, by the early 1990s, the state again became strapped for cash. With no state income tax and a patchwork tax system, revenues were expected to fall billions of dollars short of maintaining current service levels. Second, the public had become more disillusioned with state services and demanded that state government become more efficient and less costly. Legal and financial turmoil also seemed to characterize virtually all major state services. With federal government intervention threatened, public school finance (*Edgewood v. Kirby*) awaited a solution, legal and prison authorities were demanding more space for more inmates, and health and human services officials needed more monies to meet minimum legal requirements. As its comptroller lamented, Texas appeared to be "stumbling from crisis to crisis."

Performance funding in Texas evolved in only a few months.

Amid this turmoil, state policymakers reconvened in Austin. Several legislators wanted to apply appropriations based on performance to all state agencies, but particularly to higher education institutions where, for more than 25 years, funds had been allocated by formula, along with special item appropriations. THEC now was charged with developing a proposal to integrate performance funding into the formula system. In response, THEC produced in 1992 a collection of some dozen proposed performance measures assembled from other states, including minority enrollment and graduation rates, to be used in distributing a portion of state appropriations. THEC now was charged with developing a proposal to integrate performance funding into the formula system. In response, THEC produced in 1992 a collection of some dozen proposed performance measures assembled from other states, including minority enrollment and graduation rates, to be used in distributing a portion of state appropriations. When the state approached incentive funding in somewhat better times in 1987, quality and access funding was considered an add-on to state funding formulas. By the early 1990s, however, the state's public mood and dire financial situation required a more visible concern with efficiency and economy. As a result, THEC proposed a performance funding system based on a 10% deduction from existing formula levels. This aspect of the proposed system was particularly distasteful to institutions. THEC also proposed that any portion of performance funding not granted as a result of underperformance on the proposed measures be "divided among other institutions with superior performance on their selected measures."

While a comparable Tennessee system was constructed over several years, performance funding in Texas evolved in only a few months. The proposed system was, in the words of the higher education commissioner, "a work in progress." Moving into such uncharted waters rapidly and with a number of unsolved problems, such as the 10% deduction and an imprecise model, left the system without many advocates. But to publicly oppose an initiative that promised to enhance customer service and quality and to boost performance

Assessment Update
November-December 1993
Volume 5, Number 6

at a more reasonable cost was a difficult stance for higher education to take. When they did take a position, university advocates generally favored phasing the program in over a number of years, as an add-on to the formulas, and with a lower percentage of total state funding to be based on performance.

When THEC's proposal went to the legislature in 1993, it appeared initially that the system would be adopted. But in the House, the system became embroiled in a much more controversial "social engineering" proposal whereby $61 million in performance dollars would be reallocated from upper-division institutions to south Texas, predominantly minority institutions, or both. In the Senate, performance funding faced low-key opposition by higher education officials who generally resisted implementing reform in a climate of retrenchment and downsizing. Faced with growing opposition and reservations about the concept, the proposed system ultimately collapsed in the legislature. Data on the proposed performance measures will be collected from institutions by the Legislative Budget Board, but without any link to funding. The challenge was thrown out to higher education to develop a more acceptable system for future implementation.

Performance funding had promised to change radically the way Texas funded higher education, and to produce a higher plateau of service and efficiency. But the proposal faltered largely because of its use of negative incentives, because of its entanglement with racial and ethnic issues, and, above all, because it was developed so quickly. A more positive approach and better planning will be required next time if the system is to gain acceptance, and a next time remains quite likely.

A *more positive approach and better planning will be required next time if the system is to gain acceptance.*

Part H: The Shape of the "Camel's Nose"

Part H of Title IV assigns the states additional obligations for monitoring the performance of postsecondary institutions.

One of the most intriguing federal accountability developments of the past few months has occurred in an out-of-the-way place: Part H of Title IV of the recently reauthorized Higher Education Act. This component assigns the states additional obligations for monitoring the performance of postsecondary institutions. Part H is formally titled "Program Integrity Triad" and in essence requires each state to establish a State Postsecondary Review Entity (SPRE) to assume responsibility for administering an institutional review process to determine continuing institutional eligibility to receive federal funds. Not all institutions are at this point affected, of course. SPRE reviews principally will be triggered if an institution exceeds a 20% default rate on federal student loan programs, although regulations also call for reviews if an institution exhibits "excessive dependence upon funds received from the federal government," if there is a "pattern of student complaints regarding misleading or inappropriate advertising," or "other similar criteria." But because any institution could theoretically trigger such criteria, the guidelines currently call for each state to establish a general-purpose review capability. And like the proverbial camel's nose under the tent, much more could potentially follow.

Implementation of Part H is to occur in two phases. Fiscal year (FY) 1994 activities call for each state to establish a set of review standards "in consultation with institutions," as well as to initiate a mechanism for receiving and responding to student complaints. In October, 1993, each state was granted funds to cover the costs of these activities, and some (as of this writing) have already received federal approval of their plans. FY 1995 activities, which could begin as early as June, 1994, require states to begin implementing and operating the review process approved by the Feds.

In essence, Part H requires states to compile information of three quite different kinds. The first consists of information about institutional practices and procedures that can be obtained largely through an examination of published or written material like catalogues and policy manuals. Most requirements here address such issues as academic rules and regulations,

Assessment Update
January-February 1994
Volume 6, Number 1

BEST COPY AVAILABLE

records availability, and accuracy of student recruitment and other promotional material. Also in this category, though, is a requirement that each institution have "an adequate assessment process to determine students' ability to benefit from postsecondary instruction"—a provision that requires states to set specific criteria that address what such a program ought to look like. A second type of information to be mandated through each SPRE strongly resembles the "performance indicator" systems already in place in several states. This component requires assembly of such statistics as completion and graduation rates, occupational and transfer placement rates, pass rates on certification and licensure examinations, staffing patterns, and "expected returns on investment" for students in the form of a ratio between tuition and projected salaries. The final SPRE information component has a strong accreditation flavor; it covers such matters as the "appropriateness, quality, and content" of instructional programs and courses in relation to stated objectives, and the "adequacy of support" provided to these programs in the form of physical facilities, equipment, instructional materials and staff, and student support services. Together, this is a formidable array of information, and one far beyond the scope of what most states now collect.

State authorities have greeted these new responsibilities with mixed feelings. For some states (like New York), the assignment is welcomed because it legitimizes functions that state education officials have sought for some time. For others (like Tennessee), such requirements mean little because higher education coordinating bodies have been collecting and using similar data for years. But for many (including the majority of smaller and western states), Part H means a new and unwelcome staff burden.

Whatever a state's reaction, Part H is clearly part of a larger federal educational agenda that includes such components as Student Right-to-Know and stricter rules for accreditation bodies. Like the administration's more prominent approach to health care reform, moreover, this agenda involves a growing partnership between state and federal authorities to develop more efficient delivery systems and more visible accountability. Cost containment and the development of appropriate performance standards for both students and institutions, for instance, were among the items prominently mentioned as priorities for the administration in a recent address by David Longanecker, the new assistant secretary for postsecondary education (and former SHEEO in Colorado), and the proposed SPRE process was cited as an example for future state-federal partnerships. Like health care reform as well, of course, it is far too early to tell how things will turn out. But with the states now in essence mandated to get into standards-based assessment whether they like it or not, what seems certain is that our community will eventually be affected. Readers are urged to find out what is happening with the SPRE process in their own states and to get involved; judging by the early returns in this process, most states will need all the help that they can get.

This agenda involves a growing partnership between state and federal authorities to develop more efficient delivery systems and more visible accountability.

A New Look at Accreditation

Accreditation was our own, rooted in a deeply academic tradition of peer review and less tainted by the threat of alien intervention.

From the time assessment first emerged as a serious stimulus for institutional action in the mid-1980s, two external forces have worked in combination—state mandates and the requirements of institutional accreditation bodies. While fond of neither, institutions have strongly preferred the latter. Accreditation was our own, rooted in a deeply academic tradition of peer review and less tainted by the threat of alien intervention. Until recently, though, the requirements of the two processes were not very different. The vast majority of state assessment mandates that emerged in the period 1986-1990 looked a lot like what accreditors also were demanding: an institution-centered, improvement-oriented approach that emphasized the up-front specification of learning goals, allowed institutions substantial discretion in choosing assessment methods, and placed considerable value on the use of assessment results for improvement.

The strong emergence of accountability concerns at both the state and the federal levels in the last two years has changed all this. Many states have supplemented their institution-centered approaches with "performance indicators" intended specifically to fill the accountability gap; others have launched requirements for faculty workload studies, deployed performance-funding mechanisms, or have otherwise acted to address growing public concerns about higher education's lack of responsiveness. Given its structure and philosophy, this is a road accreditation cannot follow.

External attacks on the credibility of accreditation's quality-assurance role have escalated as a result. In the last reauthorization of the Higher Education Act two years ago, Congress almost eliminated voluntary accreditation from the historic triad of federal, state, and voluntary mechanisms to certify institutions to receive federal funds. The highly visible Nunn hearings on abuses of financial aid badly damaged the image of all accreditors—though cited incidents were largely from the proprietary sector. New draft rules for federal approval of accreditation bodies issued as part of the "Program Integrity Triad" (the same initiative that brought us SPREs) proved a final straw: Highly prescriptive, these regulations require accreditors to look more

Assessment Update
May-June 1994
Volume 6, Number 3

BEST COPY AVAILABLE

directly at educational outcomes (including statistical indicators on such things as job placement rates and licensure and certification pass rates), and to act as federal police in the enforcement of Title IV regulations. If voluntary accreditation is to survive as a quality-assurance player, drastic measures are called for.

In Tucson in January, 1994, directors and commission chairs from all six regional accreditation bodies met with national association representatives to agree on a new national vision for institutional accreditation. Points agreed on at this meeting include

- Creation of a permanent national organization to oversee institutional accreditation and to develop common core standards to be used by all accrediting bodies
- Establishment of a multitiered recognition system for institutions: baseline accreditation using common national standards focused principally on undergraduate education, a second tier designed to promote institutional improvement with a particular emphasis on student learning outcomes, and a possible third tier to recognize exemplary performance
- Public reporting of review results, together with confidential letters to institutions on needed improvements.

Detailed development of an implementation plan centered on these concepts was to begin immediately, with preliminary planning to be completed by the end of spring 1994.

Assessment of undergraduate outcomes *has become the principal focus of the accreditation process.*

As might be expected, initial reporting of these developments in the media has concentrated on their structural aspects: who should be in charge, the role of public membership and reporting, and so on. Much less attention has been focused on their far more radical tilt toward making the *assessment of undergraduate outcomes* the principal focus of the accreditation process. Nevertheless, the commitment to do so has been made and regardless of structure, the group agreed, this commitment remains the key to regaining public credibility for the accreditation process. If voluntary self-regulation cannot begin to meet public demands by focusing far more aggressively on what students learn and on the quality of their actual collegiate experiences, structural changes in the process will be irrelevant.

For the practice of assessment, this potential shift of emphasis has considerable consequences. While current accreditation-related assessment requirements concentrate on locally developed measures aligned with an institution's own goals, future requirements may well demand the use of more common goals and measures. And these will almost certainly center on such commonly claimed generic undergraduate outcomes as communications skills and critical thinking. What remains uncertain at this point is whether this entire initiative will be successful in preserving accreditation as an alternative to increasingly direct government assessment alternatives at all levels. Given the late date, its chances appear no better than even.

Assessment Update
May-June 1994
Volume 6, Number 3

Performance Funding: New Variations on a Theme

The idea of performance funding has intrigued state leaders for more than a decade.

Assessment Update
July-August 1994
Volume 6, Number 4

The idea of performance funding has intrigued state leaders for more than a decade, but until recently Tennessee remained its solitary example. As state "indicators" systems have proliferated, however—and as higher education's accountability problem has become more urgent—more states have started to move in this direction. Texas had a significant encounter with the concept in 1993, but largely because of its high-stakes nature this initiative remains on indefinite hold. Arkansas, Kentucky, and Missouri have proceeded less spectacularly and more deliberately.

Missouri's initiative provides a case in point. Titled "Funding for Results" (FFR), its origins date from 1989, when this state began reexamining its "base-plus" approach to public university funding and looking for a means to encourage quality improvement. As in Tennessee more than fifteen years before, state higher education leaders in Missouri were searching for new ways to "sell" higher education to legislators facing flat revenues and escalating alternative priorities for expenditure. Also as in Tennessee, FFR came out of the state's Coordinating Board for Higher Education (CBHE), which worked in consultation with the institutions themselves. Partly because of these factors, a FFR component has been successfully a part of Missouri's plan for providing funding allocations to public four-year institutions since 1991—beginning with a .5% set-aside to fund the program, with a planned increase to 2% to 3%. In 1993, this approach (both base-plus and FFR) was extended to the state's two-year colleges, whose previous state allocations were based entirely on a growth-oriented cost-per-credit-hour reimbursement formula.

Proceeding slowly, Missouri was able to learn from previous state-level performance funding initiatives and to keep up with more recent developments. In contrast to Tennessee, where outcomes measures in the form of the ACT-COMP examination and major field test results were included as performance criteria almost from the beginning, the Missouri initiative began with relatively noncontroversial and readily available statistics, including degrees awarded in identified "critical" fields and to minority students and, for two-year colleges, the number of students completing associate degree requirements or transferring to senior institutions. Only last year were outcomes measures added to FFR: for example, the percentage of students scoring above the 50th percentile on nationally normed examinations. Even these were measures already required of institutions as part of the state's five-year-old assessment initiative implemented by executive order.

The architects of FFR chose not to embed performance parameters in complex statistical formulas. Most have instead been operationalized in terms of specified additional payments to institutions in return for each "unit" of desired activity. The 1995 fiscal year allocations, for instance, award a $1,000 bonus to four-year institutions for each baccalaureate degree granted to a minority student. This renders the approach easily understandable to both institutions and the lay public. Finally, marginal dollars allocated through FFR have been kept within manageable limits; current guidelines call for FFR to be between 2% and 3% of total instructional budget, well below Tennessee's current 5.5% or Texas's politically unsuccessful 10% marginal funding proposal.

While effective in launching FFR, these features did not really break new ground. But innovation is clearly apparent in CBHE's most recent intent to add a component that enables individual institutions to set their own performance targets and to be rewarded with additional funds for attaining them. The resulting two-tiered FFR proposal was designed to overcome difficulties that have up to now bedeviled performance funding approaches. One is how to handle differences among institutions with respect to mission and student clientele without "homogenizing" performance. While retaining a component based on common state-level measures, the additional tier of FFR potentially allows institutions to be rewarded for achieving mission-specific goals. A second advantage of the approach is its potential to increase institutional involvement and ownership. By providing incentives for colleges and universities to attain targets that they themselves establish, the process encourages priority setting in the first place. At the same time, it potentially avoids a major drawback of pure performance funding, namely, that only already-strong programs and institutions are rewarded even though the real incentives for improvement may be elsewhere.

This look in performance funding affects assessment practice considerably. Like the mandates of the 1980s, it encourages the development of locally owned measures and approaches. Institutions will be unable to qualify for "tier two" funding unless they can operationalize and measure what they hope to achieve. But unlike previous mandates that often isolate assessment as "a train on its own track," the approach promises to link institutional assessment with budgeting and planning from the outset. For this reason alone, FFR's proposed second tier looks like a program to watch.

Unlike previous mandates that often isolate assessment as "a train on its own track," the approach promises to link institutional assessment with budgeting and planning from the outset.

Assessment Update
July-August 1994
Volume 6, Number 4

Acting Out State Postsecondary Reviews

SPREs are charged with developing quantified performance standards in five basic areas.

After a few late starts, all states have now established state postsecondary review entities (SPREs) as required by federal law and are fully engaged in the process of developing mandated performance standards for the institutions they are directed to review. While few states will likely be actually applying such standards until well after this column is published (only three have had their final submissions forwarded to the secretary of the U.S. Department of Education for action as of this writing), both the process they are using and the varied approaches they appear to be adopting are worth careful scrutiny because of their potential connections to future state practice in assessment.

Based on rules promulgated in 1994, SPREs are charged with developing quantified performance standards in five basic areas: graduation rates, withdrawal rates, occupational placement rates for vocational and professional programs, licensure and certification pass rates for applicable programs, and the degree to which tuition and fee charges in vocational programs are "excessive" in the light of what a graduate is likely to earn. In all five areas, moreover, the SPRE is directed not only to develop metrics but also to establish minimum thresholds of performance. All but the last of these outcomes dimensions have been the subject of considerable past discussion under the aegis of "Student Right-to-Know" (SRK), Carl Perkins, and state-level performance indicators. But the variety of ways in which states are responding to this seemingly familiar task is remarkable.

Much of the current dynamic is illustrated by the case of establishing acceptable graduation rate standards. With regard to this measure, of course, the picture is considerably complicated by the lack of final rules for SRK, a requirement enacted by Congress more than five years ago that mandated universal disclosure of graduation rates to prospective students. SPRE rules adopted in 1994 strongly suggest the use of SRK's basic cohort methodology, and a majority of states have taken the easy way out by simply referencing SRK procedures in their SPRE standards without knowing precisely what they will look like. Others, aware that proposals for SRK

Assessment Update
January-February 1995
Volume 7, Number 1

BEST COPY AVAILABLE

currently under consideration by the Department of Education depart significantly from the "first-time, full-time, fall cohort" standard for calculating graduation rates that they have applied for years to public institutions, are sticking to this procedure despite the possibility that SRK may eventually come out differently. Additional states are adopting quite different approaches, stimulated by both the lack of available cohort data on which to establish reasonable standards and the knowledge that application of cohort-based procedures will likely impact heavily on open-admissions institutions. Most prominent are methods that rely on constructing ratios between numbers of degrees granted and total numbers enrolled—in essence, a measure of "degree productivity."

More surprising, perhaps, is the variety of performance thresholds set on the graduation rate measure. About half the states have opted for a fixed minimum standard, while the balance propose an "acceptable" range beginning at one standard deviation below a comparative peer group mean. (New York, with complete unit-record data available for both public and private degree-granting institutions, has been uniquely able to put forward a regression-based model that predicts "expected" performance and then examines the degree to which a given reviewed institution lies below this prediction after taking into account a range of background factors such as admissions selectivity, size, financial aid burden, and gender composition). At the same time, most SPREs (but by no means all) have attempted to avoid holding two-year and four-year institutions to the same graduation rate standard. Actual minimum percentage rates under discussion by the SPREs for the four-year sector as of this writing, however, vary considerably: from highs in the mid-50s to lows in the mid-20s.

The visible establishment of extremely low standards may quickly catch the attention of state legislators already convinced that higher education is ducking accountability and performing ineffectively.

If it stands, this outcome may in the long run spell political trouble. Although the result of understandable institutional lobbying for the lowest possible levels, the visible establishment of extremely low standards may quickly catch the attention of state legislators already convinced that higher education is ducking accountability and performing ineffectively. In this case, as has happened before, actions against the Feds in part stirred up by higher education lobbyists in Washington appear remarkably short-sighted, often resulting in even greater unfavorable political exposure at home. At the same time, the sheer variety from state to state in what is considered "minimum" graduation rate performance is likely to yield further efforts by federal actors to "standardize" apparently irrational variation.

To be fair, nothing is final. Once the Department of Education becomes more fully aware of this emerging variation in standards as it reviews and approves each state's final submission, it may act informally to achieve greater consistency. And SPRE itself, it must be admitted, remains a shaky enterprise. Fiscal year 1994 appropriations at approximately the previous year's levels ensure the survival of the venture for at least the next 18 months—at least long enough for all states to finish the implementation process and surely long enough for some of them to conduct reviews. The mood of the new Congress is uncertain, however, and many in higher education policy circles already are broaching the question of how to reestablish the original accountability "triad" of federal, state, and voluntary accreditation on its old basis.

So might the current SPRE drama be only an episode, with little lasting impact on assessment? Not likely, I think. The bottom line is that the question of establishing acceptable minimum performance thresholds on core institutional outcomes has been raised publicly and politically for the first time in every state. And now that it's out, this topic may prove extremely difficult to put back in its box.

This topic may prove extremely difficult to put back in its box.

Assessment Update
January-February 1995
Volume 7, Number 1

PETER T. EWELL

Assessment in Tennessee: A State Profile

Tennessee's Performance Funding Program, formally initiated in 1979, remains one of the most distinctive and most often cited approaches to state-based assessment. Simple in concept, its "reward for performance" features have made it attractive to legislators in many states who are seeking ways to improve the quality, responsiveness, and accountability of public colleges and universities. To date, however, no similar programs have emerged in other states.

Because the program evolved slowly, there was time for significant campus involvement and gradual escalation of the stakes.

The roots of what was initially called the Instructional Evaluation Schedule in Tennessee go back to the mid-1970s, when the state's higher education coordinating body—the Tennessee Higher Education Commission (THEC)—faced the familiar problem of justifying new funding in the context of shrinking enrollments. The idea of setting aside a limited amount of additional money to encourage better performance was highly attractive politically, both because it supported necessary appropriations and because it linked new dollars with a tangible return on investment. In contrast to much later assessment programs in other states, the Performance Funding Program was never really seen as a method of encouraging grass-roots institutional change, but rather as a means of supporting necessary budget increases.

Because the program evolved slowly, there was time for significant campus involvement and gradual escalation of the stakes. From 1976 to 1979, the program operated as a pilot. In 1979, the legislature awarded the first funds for instructional evaluation activities in Tennessee's public institutions. By 1982, the shape of an operational program had emerged: set-aside funds would amount to 2% (later increased to 5%) of available instructional dollars statewide, and institutions would be allocated these funds on the basis of five performance criteria. Despite the *performance* label, only two of the five criteria—"value added" on the newly developed ACT-COMP examination, and major field testing with standardized exams—relied on actual test results. Two more—accrediting all accreditable programs, and surveying students on their satisfaction—rewarded institutions just for

Assessment Update
Fall 1990
Volume 2, Number 3

BEST COPY AVAILABLE

engaging in the assessment process. The fifth criterion was based on the degree to which institutions actually used this information to make local improvements.

Despite initial doubts and opposition, most institutions adapted reasonably quickly to the new program. Some, including the University of Tennessee at Knoxville, made effective use of the leverage it provided to significantly enhance their own local assessment, planning, and program review activities. By the end of the initial five-year authorization period, however, many were "topping out" on the process-oriented criteria, and THEC moved to realize the program's original performance-based design.

New performance funding guidelines, issued in 1988 and currently in use, emphasize assessment results. For general education, the ACT-COMP has been retained, although institutions now earn performance funding dollars both for score gains and for the absolute level of scores obtained. Dollar allocations in major fields are now based on actual test performance. Moreover, to ensure equitable performance judgments in the area of alumni satisfaction, a set of common statewide survey questions was developed, and results are used to allocate funds

Performance funding in Tennessee is seen as a successful venture.

The merits and drawbacks of these changes, widely debated in Tennessee, are instructive for all other states and institutions. On the one hand, the new guidelines are considerably less ambiguous in their application, an important consideration when dollars are at stake. On the other hand, institutional observers see score-based performance criteria as inherently antithetical to developing the kind of faculty involvement most likely to produce instructional improvement.

On balance, however, performance funding in Tennessee is seen as a successful venture. From the THEC perspective, a major benefit is enhanced ability to talk effectively to the legislature. "It's very impressive when you don't just lead with a funding request," emphasizes the current THEC executive director. "The presence of these dollars helps us sell the rest of the budget." At the institutional level, moreover, the program has clearly provided incentives to develop a local assessment process that would otherwise not have gotten off the ground.

But both constituencies can also cite some unanticipated costs. Legislators concerned about higher education increasingly recognize that the logic of rewarding success, while compelling, may direct money away from the problems where it is most needed. In part, this motivation underlies a newer "legislative goals" initiative in Tennessee, whose purpose is less tied to rewarding performance. Institutional leaders have found that the existing performance funding criteria may unintentionally be strengthening the hand of "sectional" curricular interests, particularly where national exams are available or where professional accreditation is at stake. As one campus administrator summarized the situation, "I'm worried about institutions feeling constrained by performance funding into doing things that lack local utility . . . Maybe in this case you shouldn't go after these funds." In part, it

Assessment Update
Fall 1990
Volume 2, Number 3

is these attributes that have deterred other states from emulating Tennessee's assessment experiment.

Nevertheless, many of us continue to learn a great deal from Tennessee. As part of a pioneering venture, Tennessee's assessment efforts have been unusually well documented. As a result, they have effectively informed much of current assessment practice at both state and institutional levels.

Many of us continue to learn a great deal from Tennessee.

Assessment Update
Fall 1990
Volume 2, Number 3

So, Are They Really Going Away?

Many believe that higher education has "won" the accountability battle and that mandated assessment will wither as a result.

In the wake of last November's election, national accountability initiatives seem to be melting away as quickly as they first appeared. Many of the provisions of Part H of the reauthorized Higher Education Act—including the highly controversial State Postsecondary Review Entities (SPREs)—are in deep trouble. Other federal reporting requirements such as Student Right-to-Know remain unregulated. And attempts to refashion accreditation on a national basis, begun last year by the higher education community in reaction to what was then seen as an unprecedented federal regulatory onslaught, have been largely abandoned after encountering massive resistance in the field.

At the state level, meanwhile, big government appears also to be on the run. An extraordinary number of states are decentralizing their governance systems, emphasizing local control while downsizing (or eliminating) state higher education agencies. Among the casualties last fall was the New Jersey Department of Higher Education, once prominent as the initiator of the most comprehensive state approach to accountability ever launched (see *Assessment Update*, 1989, Vol. 1, No. 2).

Viewing these developments, many believe that higher education has "won" the accountability battle and that mandated assessment will wither as a result. But before the dancing in the streets begins, it is important to look a little deeper at what has been going on.

At the federal level, three themes are apparent in recent rhetoric, none of which seems likely to remove assessment permanently from the policy table. *Regulatory relief* is a central theme, and the one most responsible for the back-off on SPRE and other federal oversight initiatives. But behind this slogan is also the conservative notion of protecting entrepreneurship. Historically, deregulation by Republicans has applied to the tax-paying public, not to the recipients of public funds. On the receiving end, its counterpart is "paying for results." This notion has already been rehearsed at the state level in new performance-funding schemes in states such as

Assessment Update
July-August 1995
Volume 7, Number 4

BEST COPY AVAILABLE

Missouri, Kentucky, and Arkansas. It's also visible in an approach floated in April 1995 by the Department of Education (DOE) that proposes easing up on Title IV oversight for institutions that can show they meet a set of minimum statistical performance thresholds—including outcomes.

A second prominent theme in Washington is *devolution* of authority (and funding) to the states. Early congressional proposals to consolidate DOE programs also emphasize the use of block grant mechanisms and other forms of state-federal partnership to administer this more "streamlined" program array. But such an approach also has accountability implications, as states are presumed to know best both how to gatekeep institutional access to funds and assess the effectiveness of such programs. Given this policy direction, something that looks like a SPRE appears inevitable.

A third theme is one of better focusing public programs to produce more clearly defined (and largely utilitarian) public benefits. For postsecondary education, increasingly this means work force training. Again, the dust has yet to settle, but a number of bills introduced in Congress this spring attempted to consolidate a range of DOE, Labor, and other federal programs that affect vocational education into a single package. These proposals echo recent rhetoric at the state level about the proper direction of postsecondary investment, as well as the Clinton administration's own support of "school-to-work" initiatives. With ends such as these, the assessment implication is clear: a growing center of accountability will be what happens after college, not inside it. And this is a prominent feature of current SPRE performance standards, as well as a growing number of state-level performance measures.

A growing center of accountability will be what happens after college, not inside it.

These themes together do not suggest a diminished role for assessment as an element of accountability in the long run, though they may well signal the demise of a largely aberrant episode of federal oversight. The 1992 Amendments to the Higher Education Act served as a blunt instrument—intrusive, duplicative of existing authority, and badly crafted to boot. Student Right-to-Know, passed in an earlier fit of congressional pique, had many of the same defects. But both have had considerable indirect impact in the precedents they set for the states. Some 18 states now regularly report graduation rates for public institutions using methodologies originally developed to meet never-finalized Right-to-Know requirements. SPRE performance thresholds, now that they are developed, are similarly ready-made for later state-level application, whatever the ultimate fate of SPRE as a federal program.

In light of these secondary effects, calling federal accountability efforts ineffective may be badly off the mark. Three eventful years of federal follies are ending up largely returning the accountability ball to the state court where it started ten years ago. Partly because of the ensuing game, the states now appear far more prepared to play it. As rhetoric turns to business in the coming year, we should not be surprised if they apply these newfound skills.

Assessment Up Front

A growing number of states are developing new approaches to assessing students for admission to their public colleges and universities.

For a number of years now lawmakers have been criticizing higher education for its apparent lack of engagement with the massive changes in assessment and pedagogy taking place in the nation's public schools. Admittedly, these sentiments have a basis in fact: Policymakers see a large cohort of students proceeding through the K-12 system who have experienced authentic modes of testing and active modes of teaching that colleges and universities are ill-prepared to receive. But in assessment discussions especially, the issue has in part been symbolic: how can higher education institutions seriously claim that the outcomes of what they do are complex and "unmeasurable" while continuing to accept uncritically such measures as standardized test scores and courses completed as adequate bases on which to assess prior achievement?

Spurred by such concerns, a growing number of states are developing new approaches to assessing students for admission to their public colleges and universities. Oregon's Proficiency-Based Admissions Standards Study (PASS) program, most observers would agree, has been the pacesetter in this arena. This effort can trace its immediate ancestry to legislative action in 1994 in the form of HB 3565—a measure intended explicitly to "enhance alignment" between the reforms occurring in the state's K-12 system and practices in higher education. Already being implemented at that time in Oregon (as in other states) was a reengineered high school exit assessment based on "authentic" techniques consistent with the nationwide "New Standards" Project, and concerns were increasingly being expressed that this would render current course-based collegiate admissions requirements both inappropriate and obsolescent. Responding to these concerns—made concrete by HB 3565—the state's system of higher education officially adopted the PASS program in 1994, with the aim of completely replacing current collegiate admissions requirements with new assessment-based proficiency standards by 1999.

The actual implementation of this ambitious effort is decentralized, based principally on formal working partnerships among the principal parties-at-

Assessment Update
January-February 1996
Volume 8, Number 1

BEST COPY AVAILABLE

interest: high schools, community colleges, and state universities. Four such partnerships were established in 1995, and sixteen are currently in place. Each group is charged with developing and piloting assessment approaches centered on six content areas: math, sciences, social sciences, foreign languages, humanities and literature, and fine and performing arts. Each content area is governed by commonly adopted statewide competence statements that describe the abilities being assessed. In addition, these performance assessments are to evaluate students on an additional nine "process proficiency areas," including reading, writing, oral communications, critical and analytical thinking, problem solving, technology, systems and integrative thinking, teamwork, and "quality work." The specific assessment technologies being developed under this rubric are intended to do double duty, both certifying mastery at high school exit and serving to admit students to college. While colleges and universities would remain free to use Scholastic Aptitude Test and American College Test scores as supplements, the resulting system is intended to replace altogether the current official standards, which are based on prior completion of a prescribed college preparatory curriculum.

Other states following this path are proceeding more deliberately. For instance, in Kentucky, where K-12 reform has also been in full swing for a number of years, current course-based admissions requirements remain in place. But a parallel effort under the aegis of the state's Commission on Higher Education, titled the "Kentucky Instructional Results Information System" (KIRIS), is piloting additional assessment alternatives that strongly resemble those being implemented in Oregon, including both portfolios and specially designed extended-response production tasks. Study commissions in states such as Florida and New York are following similar paths.

Assessment actually takes root on a campus when it is approached not as an accountability add-on but as an integral part of how business is done.

As PASS proponents in Oregon clearly recognize, the implications of such efforts for assessment's future place in higher education are inescapable, controversial, and profound. Governing traffic at a key "transition point" in the education continuum on the basis of assessed competencies rather than on traditional Carnegie-based measures of exposure (which a recent Florida report aptly labeled "butt-time units") raises the potent question of why such approaches cannot equally appropriately be used to certify mastery within college or in transferring from one college to another. Coupled with the increasing prominence of instructional technology—which also threatens to explode current course-based units of academic accounting by allowing students to master intended competencies at different times in alternative ways—the result is a potentially powerful impetus to remake college curricula in ways that embed assessment at every stage.

For more than a decade, pioneers at Alverno have reminded us that this is how assessment actually takes root on a campus—when it is approached not as an accountability add-on but as an integral part of how business is done. As an increasing number of states begin deploying serious, performance-based assessments at collegiate "front doors," it will be instructive to see how, and whether, they can exploit the implied opportunity to slip inside.

Assessment Update
January-February 1996
Volume 8, Number 1

On the MARC: The Efforts, Obstacles, and Successes of the Maryland Assessment Resource Center

> ***The center was not intended to be a permanent fixture but rather a support service for the initial efforts of institutions to comply with the state assessment and accountability mandate.***

Guest Columnists: Robert W. Lissitz, chair and professor, Department of Measurement Statistics and Evaluation, University of Maryland, College Park; and Maria Wainer-Yaffe, assistant to the dean, College of Education, University of Maryland, College Park.

In 1988 the Maryland General Assembly passed the Higher Education Reorganization Act, also known as the Performance Accountability Law, with the aim of making student outcomes information available to legislators and citizens. In 1990 the Maryland Higher Education Commission (MHEC) adopted a policy requiring the governing boards of public colleges and universities to submit campus assessment plans and to report annually on progress in implementing those plans.

This article summarizes Maryland's approach for providing assessment support through the creation of the Maryland Assessment Resource Center (MARC) housed at the University of Maryland at College Park. Funded by MHEC, MARC operated from August 1991 to July 1995 with the aim of supporting efforts at all two- and four-year public institutions of higher education throughout Maryland. A twelve-member advisory committee, including vice presidents, directors, coordinators, and professors from community colleges and universities throughout Maryland, as well as MHEC and University of Maryland administration, established guidelines for the center's operations.

Assessment Support Efforts

MARC was designed to serve as a catalyst for the assessment efforts of 31 two- and four-year public institutions of higher education throughout Maryland. The center was not intended to be a permanent fixture but rather a support service for the initial efforts of institutions to comply with the state assessment and accountability mandate. As with most new ventures, and consistent with the research literature, we found that there were hurdles to

Assessment Update
May-June 1996
Volume 8, Number 3

BEST COPY AVAILABLE

overcome, including an absence of allocated funding to conduct studies, a lack of faculty experience in conducting assessment studies, and resentment toward the state mandate. These obstacles were addressed by MARC with a variety of approaches directed toward faculty and administrators, including (1) conferences and workshops, (2) an annual MARC Outstanding Assessment Award, (3) a newsletter designed to disseminate information on assessment projects on Maryland campuses, (4) the volunteer Technical Support Group composed of academics willing to provide free consultative services and review grant proposals, (5) a matching grant program that awarded funding for campus assessment activities involving faculty, and (6) research projects.

MARC conducted two research projects. One, analyzing reviewer feedback on grant proposals, was aimed at determining the problematic areas in proposal submission. Findings served as the basis for a MARC newsletter article and a segment of a MARC workshop designed to make faculty assessors aware of methodological pitfalls.

The second research project was a survey designed to determine the extent to which grant-funded studies were being utilized. Findings revealed that a number of studies took longer than project coordinators had anticipated. Changes that had been implemented included curriculum revisions; program expansion; process and policy revisions; standards, mission, and goal redefinition; new awareness and cooperation; and data collection and monitoring efforts. The most frequently cited changes involved perceptions. These included faculty awareness and cooperation and focused on efforts to inform faculty through publications, videos, training sessions, and workshops. The major obstacles to making changes were financial and social.

Those who overcame financial obstacles either obtained additional funding or proceeded more slowly. A gradual process was advocated by some who were confronted with resistance. Administrative support and involvement of others were considered most important in overcoming the social obstacle. Respondents offered advice, including realizing that the project may take longer than planned, carefully planning budget estimates, using pilot-testing, following up on the impact of workshops, documenting details for reallocation efforts, and narrowing the field of research. The importance of communicating and support building was the most frequently given advice. Garnering support initially and throughout the process, persuading administrators who control funding, and demonstrating that changes will help faculty with their own tasks and not add work were all critical messages to convey.

The most disturbing finding of the study was the prevalence of mixed messages delivered by institutions with regard to the importance of assessment.

The most disturbing finding of the study was the prevalence of mixed messages delivered by institutions with regard to the importance of assessment. Academics expressed frustration over lack of financial support, lack of faculty release time, lack of recognition of faculty assessment efforts, resistance of colleagues, and lack of administrative support. They felt it was not reasonable to expect faculty assessment efforts to prosper in the face of such obstacles.

Respondents assessed MARC as being very supportive of their assessment efforts. In addition to the financial aspect, academics cited MARC as enlightening the academic community on assessment issues, serving as a communication link among Maryland institutions, and providing external rewards, recognition, and support. A number of respondents expressed regret at the closing of the center.

Limitations

MARC's efforts and this study both experienced limitations. It must be noted that the respondents to the MARC survey and the participants in MARC's programs were not necessarily representative of Maryland's academics. Although the public may demand accountability and the MHEC may mandate assessment, there was no requirement for academics to apply for MARC grants or participate in MARC conferences. Those who interacted with MARC had at least some interest or had experienced political pressure to comply.

Probably the greatest obstacle faced by MARC was the human factor.

Another challenge faced by MARC was addressing the methodological needs of academics. Those already interested in assessment could be assisted by the comments of grant reviewers or the advice of a member of the volunteer Technical Support Group. However, academics not especially interested in assessment were more hesitant to pursue assessment projects since they lacked methodological awareness.

Probably the greatest obstacle faced by MARC was the same as that encountered by so many of the project coordinators, the human factor. Resentment toward the state mandate, resistance to change, and lack of interest by administrators and other personnel hindered the center's efforts to encourage assessment. Administrators' lack of interest in change or unwillingness to utilize assessment study findings sent a contradictory signal to those academics who were implementing assessment activities.

Need for an Assessment Center

A state mandate for academic accountability is not, in itself, an effective means to encourage campus assessment studies. Such a requirement alone does not address academics' concerns regarding financial allocations, methodological education, personal and political reactions, and the value of assessment. Issuing an order without providing the means to overcome these obstacles is a hollow initiative. Creation of an assessment support service more genuinely demonstrated the state's regard for the importance of bringing about change in campus accountability efforts. The center provided assessment education, established a network to facilitate awareness, served as a resource for those requiring guidance, reviewed and gave advice on early assessment efforts, created a positive image of the value of assessment, and granted funds for assessment studies. MHEC helped do all this through its funding.

Assessment Update
May-June 1996
Volume 8, Number 3

Conclusion

An assessment support service can offset some of the obstacles to implementing a state mandate by offering funding and efforts designed to increase awareness. However, an assessment center makes an additional contribution to the process of institutionalizing assessment by providing a period of adjustment for all parties. Institutions, administrators, and faculty can make themselves aware of assessment issues, establish connections, commit financial allocations, create policies that support faculty assessment efforts, and develop an assessment plan during this period. Theoretically, when the dissolution of such a center comes about, academe will be prepared to incorporate assessment at all levels of the enterprise.

The center made a significant difference to the assessment environment in the state of Maryland. It is verifiable that scores of assessment studies were conducted that would not have been were it not for MARC's support, interinstitutional contacts were created due to MARC's efforts, academics were educated on assessment issues, and project coordinators were granted prestige for their assessment efforts. Hundreds of students and other campus constituents may yet be touched by the impact of the utilization of assessment study findings. If the extent to which MARC will be missed by members is an indication of its success, then, judging from the comments on the member survey, MARC reached an impressive level of accomplishment.

The center made a significant difference to the assessment environment in the state of Maryland.

Speaking Virtually: Assessment and the Western Governors' New University

This institution is intended, through assessment, to independently certify and credential learning gained chiefly through the electronic media.

Assessment Update
September-October 1996
Volume 8, Number 5

On more than one occasion in this column, I've noted that assessment activities are gradually entering the "mainstream" of state-level attention by being embedded as essential components in far more comprehensive initiatives. Competence-based admissions experiments like those in Oregon, Florida, and Wisconsin (see this column in *Assessment Update*, Vol. 8, No. 1) and continuing discussions about restructuring state-level articulation policies provide one prominent set of examples. But the latest and most far-reaching instance—as well as the most controversial—is the action taken this past June by ten governors to create a "virtual university" to serve the western region. Officially titled the Western Governors' University (WGU), this institution is intended, through assessment, to independently certify and credential learning gained chiefly through the electronic media, including real-time video linkages and the Internet.

The origins of this initiative lie in two related phenomena that have become increasingly salient for the Rocky Mountain West. The first is the familiar call to accommodate increases in new student demand within a flat resource base, and to provide a full range of programs to sparsely populated regions. Championed most visibly by Utah's Governor Mike Leavitt, increased use of distance-education technology is seen as one way to address the resulting productivity dilemma. Another motivation is the frequently expressed goal of other governors, including, most prominently, Colorado's Roy Romer: stimulating higher education to adopt clear standards of graduate achievement and to credential students through demonstrations of achievement instead of credits based on "seat time." A university without walls, if actually implemented, would provide an unparalleled opportunity to further both agendas on a very large scale.

In basic concept, the idea is not new. Institutions such as Regents College of the University of the State of New York (and its counterpart, Thomas Edison in New Jersey) have for many years awarded degrees on the basis of testing. WGU would build on this foundation by developing a range of independently administered assessments that match collegiate-level credentials that are in particularly high demand and that are capable of "capstoning" courseware and other instructional materials already available to the region piecemeal through the electronic media. In the early stages, these credentials are likely to populate associate-level degree programs and a variety of industry-based occupational and professional certifications. Other functions of the institution are designed to ensure that potential students have access to these materials and that the resulting "programs" meet state need. To accomplish the former, WGU will create an electronic catalogue, accessible to students over the Internet, that contains entries describing particular electronic courses and learning experiences offered by providers located throughout the region. To address the latter, WGU will establish a request-for-proposals process to solicit new courseware in high-demand subject areas.

A credible set of assessments is, of course, the key to making all of this happen, and ideas about how these assessments will be designed and implemented are just beginning to emerge. Among the working principles now being used by WGU planners to design the required system are the following:

A credible set of assessments is, of course, the key to making all of this happen.

Developmental. The assessment approaches used will be built to meet diagnostic as well as certification functions, as part of an ongoing advisement and mentoring process.

Nonexclusionary. The assessments themselves, and access to the resulting credentials, will be open to all comers regardless of previous higher education experience.

Nonpunitive. Wherever possible, students will be required to retake only those portions of an assessment that they have not previously mastered—unless simultaneous demonstration is needed to ensure that individuals possess an integrated set of practice skills.

Portable. The credentialing process will be designed to provide assurances of abilities that speak to a broad range of potential clients, including specific employers and educational institutions.

Performance-based. The assessment processes used for credentialing will be based substantially on direct demonstrations of the abilities in question, even if this means relaxing the purely "virtual" nature of WGU. (Performance-based assessments associated with specific degrees or certificates will probably be administered at "franchised" local service centers distributed throughout the region.)

Accomplishments as well as abilities. The credentialing process will involve documenting past achievements and abilities—many of which are expected to be work- or service-related—probably by means of a portfolio.

Technology-intensive. Intensive use of electronic media will be built into WGU's assessment function from the outset, including the use of computer-adaptive, distance-delivered assessments and media-intensive simulation problems and demonstrations. It is generally recognized, though, that adequate demonstrations of many of the abilities to be assessed will require direct observation.

Rather than developing the bulk of the required assessments itself, WGU will likely contract the job to established testmakers, all of whom have evidenced considerable interest in the initiative. WGU would provide an excellent setting for applying new assessment technologies that testmakers have been developing.

As I write, it is too early to tell how WGU will fare. Political opposition from established institutions (many of whom have already staked out a considerable distance-education business) is growing, and many details of implementation have yet to be worked out. For example, will or should WGU be accredited and by whom? Can WGU charge in-state tuition for out-of-state students? Start-up costs are high ($10 million, by the planners' estimates). Despite these obvious drawbacks, the power of the virtual university concept among key political constituencies is apparent. And much of the appeal rests on the fact that the virtual university puts assessment visibly at the operation's center as an alternative to traditional credit-based academic accounting. Whatever WGU's fate, we should not be surprised to see this core feature surface in other guises.

WGU would provide an excellent setting for applying new assessment technologies that testmakers have been developing.

Assessment Update
September-October 1996
Volume 8, Number 5

PETER T. EWELL

Putting It All on the Line: South Carolina's Performance Funding Initiative

When it comes to mandating accountability for higher education, South Carolina's legislature has pretty much led the pack. When institutional effectiveness provisions were written into legislation on broad-based education reform passed in 1988 ("the cutting edge"), South Carolina joined a handful of states requiring all public institutions to engage in assessment. With Act 255 in 1992, the state was also one of the first to go the next step—requiring comparative performance measures for public colleges and universities in report card format. In spring 1996, the legislature was at it again, passing Act 359, which requires that 100% of all state funding for higher education be allocated to colleges and universities on the basis of formally established performance measures.

When it comes to mandating accountability for higher education, South Carolina's legislature has pretty much led the pack.

In addition to refining mission areas for public institutions, Act 359 specifically delineates some 37 performance measures that are intended to replace the state's enrollment-driven funding formula by June 1999. Grouped under nine subheadings, the list itself contains few surprises. Prominent within it are such items as graduation rates, the accreditation status of accreditable programs, faculty workload measures, and measures of general efficiency (especially in administrative operations—all of which have figured prominently in the performance measure initiatives of other states. But also apparent are interesting attempts to influence institutional good practice, an intent signaled earlier in South Carolina by Act 255's inclusion of "numbers of undergraduates participating in faculty research" as one of eleven report card indicators. Act 359 expands this category by including such domains as posttenure review, the availability of faculty to students outside the classroom, use of technology, and collaboration and cooperation across institutions and with the private sector. Such matters delve deeply into the ways institutions do business. At the same time, they are unusually resistant to the development of simple statistical measures.

Since July 1996, a steering committee and three task forces have been working under the auspices of the state's Commission on Higher Education to meet a January 1997 submission deadline for an implementation plan.

Assessment Update
January-February 1997
Volume 9, Number 1

BEST COPY AVAILABLE

By October, this effort had yielded relatively detailed operational definitions for each mandated measure, and, at present, sectoral committees are considering specific performance benchmarks and weights to drive an allocation formula. This represents an enormous achievement in a very short period of time (tight timelines are also characteristic of South Carolina's approach to accountability mandates), but the effort also has led the state in interesting directions.

First, it is not surprising that a lot of previous work on developing measures of institutional effectiveness has been incorporated wholesale. Much of this, of course, came from Act 255 and a good deal of it from the Southern Association of Colleges and Schools (SACS) and the Southern Regional Education Board (SREB) as well. Measures of faculty qualifications, for instance, draw heavily on established SACS criteria, which are far more explicit on this matter than those of other regional accreditation agencies. Measures of faculty compensation and institutional spending patterns, in turn, heavily reference benchmarks for peer states already established by the SREB. Workable approaches drawn from more unlikely sources are also being dusted off; in both job placement and graduation rate discussions, for example, traces are visible of the commission's substantial prior investment in developing outcomes measures for the State Postsecondary Review Entity. At least as important, South Carolina's long-term commitment to fostering an assessment infrastructure is paying off. Institutional assessment practitioners—through the South Carolina Higher Education Association Network originally established with state support in 1988 to develop institutional effectiveness guidelines—are reluctantly but prominently involved in the effort to design a workable system.

South Carolina's long-term commitment to fostering an assessment infrastructure is paying off.

Despite these advantages, the task remains daunting. Although at least seven states now have performance funding schemes in place that incorporate outcomes to some degree, all rely on marginal incentives amounting to no more than about 5% of institutional allocations. Two previous attempts to base all state support on performance measures—one in Florida last year and another in Texas two years ago—were tabled as politically unworkable (though both remain formally in place). It is therefore not surprising that practitioners in South Carolina are trying hard to limit the system's scope. One such attempt was a proposal to put most of the weight in the performance formula on its mission focus criteria—essentially, whether the institution has a mission statement consistent with state guidelines and a program array to match. If successful, this approach would have effectively constrained the amount of money affected by harder performance criteria to about 10% to 15%. Although this initiative was quickly ruled out of order by the commission, more recent findings suggest, at least for the moment, that the state will follow the more traditional model of establishing a separate (though substantial) "performance pot" of resources to be allocated in addition to base.

Assessment Update
January-February 1997
Volume 9, Number 1

Adding to the challenge is the fact that 37 indicators are a lot. Ironically, this could add a measure of built-in stability, if only because statistical fluctuations in a large number of equally low weighted dimensions will likely buffer any large changes in the bottom line. More seriously, many of the proposed performance criteria overlap in scope, and some are clearly more measurable than others. Mission focus measures, for instance, begin with relatively straightforward examinations of program structure, but they also raise questions about cost-effectiveness and goal achievement that are covered later by other, much better-conceived, measures. While a short-term solution to this problem has been found by allowing sectors to zero-weight redundant subdimensions to remove them from the formula, technical amendments may eventually be needed to clean up the system.

South Carolina's experience illustrates the difficulty of writing performance measures directly into law—a problem avoided by more mature applications of the performance funding concept in states like Tennessee and Missouri. At the same time, the many measures intended to tap institutional good practice remain essentially judgment calls. The proposed indicator on post-tenure review of faculty, for instance, tests the institution's adherence to a set of a dozen best-practice criteria for such processes, which will require somebody to review each institution's practices against this list to determine performance. Who will do this job and how they will behave when they know that such judgments will actually move money around remain to be seen. Again, prior experience in such states as Tennessee is instructive: over time, allocation criteria have tended to rely more and more on hard (and increasingly refined) statistical measures in an attempt to head off fix charges of bias and to remove the inevitable political bargaining associated with judgment-based measures.

At least six states have performance funding initiatives on the table for spring 1997.

Whatever its eventual shape, South Carolina's program is probably a portent of things to come. Act 255 was widely imitated by legislatures in the early 1990s as states rushed to adopt report card measures, and at least six states have performance funding initiatives on the table for spring 1997. Similarly, Act 359 illustrates virtually every policy and technical dilemma associated with such an initiative. So whatever the outcome, other states will have something to learn. For both reasons, the South Carolina case will be a fruitful one to watch in the coming year.

New Looks in Accreditation

When the accountability-for-results movement swept the country in the mid 1980s, regional accreditors had a lot of catching up to do. While several—most notably, the Southern Association and the North Central Association—gingerly evolved outcomes-based criteria in this period, the clear pressure on institutions to get started in assessment came from the states. But in the past five years this center of gravity has changed. About as many institutions now report that their assessment efforts are being developed in response to accreditation demands as in addressing state mandates.

About as many institutions now report that their assessment efforts are being developed in response to accreditation demands as in addressing state mandates.

Partly this is because some of the steam has gone out of state-based assessment. Racked by fiscal pressures and competing demands for policy attention, a number of states have effectively put assessment on the back burner; mandates remain on the books, but they are rarely enforced with vigor. Responding to the same set of stimuli, other states have evolved far more proactive approaches to accountability—including performance measures and performance funding—that are markedly different from the institution-centered, assessment-of-learning initiatives that appeared in the 1980s. Accreditation bodies, meanwhile, have moved steadily to implement institution-centered approaches. As the reauthorization process for the Higher Education Act unfolds this year, moreover, accreditors realize that they are under the gun to demonstrate their ability to play a credible quality assurance role lest they be shut out of the process, as almost happened in 1992. Two quite different examples of this increasing vitality are worth noting, the first showing the kinds of creativity individual accreditation commissions are exhibiting and the second illustrating their growing willingness to work together.

The first story concerns recent attempts by the Western Association of Schools and Colleges (WASC) Senior College Commission to design a new approach to self-study based on a combination of special topics reports and statistical data portfolios. Several accreditors, of course, have already experimented with topical approaches to institutional self-study in which qualifying institutions can focus on a few issues of their choice instead of preparing the usual

Assessment Update
July-August 1997
Volume 9, Number 4

BEST COPY AVAILABLE

standards-centered report. But WASC's experiment has a new wrinkle because it proposes to displace the usual standards-based compliance report entirely with a combination of descriptive and performance statistics packaged as exhibits in a single institutional portfolio (accompanied by certified stipulations of procedural compliance with noncountable items such as conducting regular financial audits and having human subjects research policy). The approach is being piloted with several University of California (UC) campuses and the University of Southern California (USC), institutions that are pretty safe from a reaffirmation standpoint, whose database and research capacities are well above average, and that have long complained that traditional standards-based self-studies add no value to them and are a lot of work to boot. Guidelines for the process allow considerable variation in the contents of data portfolios. But in preparing them, institutions must conform to the spirit of the standards, including paying visible attention to undergraduate learning and development of a culture of evidence. One virtue of this approach is that it allows institutions to capitalize on existing databases and assessment studies that have a visible connection to their own internal processes and priorities. At USC and UC Santa Barbara, for instance, portfolios are being constructed around each institution's own strategic plan; at UC Berkeley, a system of effectiveness indicators for undergraduate education that had already been designed for internal monitoring and evaluation purposes can be incorporated directly into the portfolio. A second, as yet unrealized, potential of this approach is in external communication. A limited number of well-chosen strategic indicators, extracted from each institution's portfolio, could eventually enable WASC to provide external audiences with an easily digestible statistical portrait of the universities in its region.

WASC's experiment has a new wrinkle because it proposes to displace the usual standards-based compliance report entirely with a combination of descriptive and performance statistics.

The second interesting accreditation story is about the alignment of standards across regions. This matter was a centerpiece of the short-lived National Policy Board (NPB) agenda, adopted by accreditation leaders in the wake of the highly intrusive 1992 amendments. The NPB effort was quickly abandoned in 1994-95 when it became apparent that institutional opposition to it was considerable and that federal interest in pursuing centrally imposed quality assurance was fading fast. But accreditation leaders recognized continuing needs both to get their own collective act together and to work out an appropriate division of labor among the states, the federal government, and themselves. At the same time, they acknowledged that new technologies for instructional delivery, as well as new patterns of student behavior, were rapidly making the regional basis of accreditation problematic in any case.

To help address the first concern, representatives of the states, the U.S. Department of Education, and the regional accreditation bodies met last spring under the auspices of the Education Commission of the States to discuss how their collective quality assurance task might best be approached in the current deregulatory policy environment. One result of this ongoing discussion appears to be a commitment on all sides to increased data sharing to detect and deal with problem cases of fraud and abuse. All three parties also are interested in placing greater emphasis on the use of common data definitions and indicators of institutional effectiveness. "Leading with information" thus appears to be the emerging policy answer to the dilemma of

Assessment Update
July-August 1997
Volume 9, Number 4

how to discharge needed quality assurance functions in the absence of direct regulatory control.

A more immediate stimulus to collective action is the rapid emergence of institutions based on distance delivery, which by definition violate established regional boundaries. The most prominent example so far is the Western Governors University (WGU), now involving fourteen states, which promises to grant competence-based degrees and credentials to students engaging in a wide range of educational experiences. Signatories to the WGU accord thus far are located in four distinct accreditation regions—North Central, Northwest, and the WASC Senior and Junior College Commissions. In negotiation with WGU, all four commissions have agreed to act in concert to review the institution. This approach avoids the need for WGU to approach each agency individually, and perhaps get embarrassingly different answers. More important, the affected commissions know that WGU is only the first of many such cases that they will be facing in the future. The most immediate result of this new collaboration was the establishment of the Inter-Regional Accrediting Committee, which will meet in summer 1997 to establish a common set of eligibility criteria for WGU. Somewhat ironically, the first step in this process involved resurrecting common eligibility criteria originally prepared by commission staffs in support of the failed NPB effort two years ago.

This surge of interest in performance among regional accreditors seems to be good news for assessment.

The potential implications for assessment of this new vitality are considerable. On the one hand, accreditors are eager to reassert their historical claim as the guarantors of academic quality amid continuing political doubts about their ability to make really tough decisions. In several cases (and perhaps most prominently in the Southern Association region these days), this claim is leading to unusually zealous enforcement of institutional effectiveness standards. On the other hand, accreditors are extremely interested in evolving new approaches that respond to institutional demands to streamline the accreditation process, and at the same time producing results that communicate succinctly to the public. One probable convergence point of these quite different imperatives is increased emphasis on institutional performance statistics—as in the WASC data portfolio experiment. This would echo the state response to similar conditions five or six years ago, when the first report cards on public colleges and universities appeared. As such data become more prominent as the currency of accountability, a longer-range implication will be the need to certify institution-generated performance statistics in much the same fashion as independent financial audits now provide public credibility for fiscal bottom lines. Indeed, some institutions already engage in third-party audits of the data they give to public information purveyors like *Peterson's*, *Barron's*, or *U.S. News and World Report*. Whether these developments occur, of course, remains open to speculation. Whatever happens, though, this surge of interest in performance among regional accreditors seems to be good news for assessment, if only as a counterweight to an increasingly aggressive and capricious round of government mandates.

Assessment Update
July-August 1997
Volume 9, Number 4

PETER T. EWELL

Technology Facilitates National Data Collection Strategies

Since the State Postsecondary Review Entity initiative collapsed in the wake of the 1994 congressional elections, there has been little public talk about accountability-related data collection for higher education at the national level. A largely deregulatory policy environment, coupled with pressures to cut back federal investments in data collection to help achieve balanced-budget goals, have pretty much ended conversations about obtrusive, big-ticket items like national tests or federally mandated measures of institutional performance. Meanwhile, though, new conversations about the future shape of national data collection are emerging that raise both interesting and unsettling possibilities. And like much that's being stirred up these days, the principal culprit is technology.

At one level, enormous increases in computing capacity, together with major strides in the technology of record linking and analysis, are allowing databases of many kinds to be tapped on a national basis. Foremost among them at this point are the unemployment insurance wage records maintained by each state, which contain unit-record information on earnings and employment for all employees in the state. Such states as Florida, Texas, Maryland, and Washington have been especially active in using these records to construct performance statistics on former-student employment, job placement, and "return on investment"—without having to rely on costly follow-up surveys, which often experience less than satisfactory rates of response. Pooling data across states and examining additional federal employment databases to identify individuals in the military and in federal civil service positions can increase the coverage of this approach substantially. Pioneering states are also linking student records in other dimensions—for instance, driver's license, prison, and welfare records—to conduct various kinds of outcomes studies. Virtually all states now have in place or are constructing unit-record enrollment databases capable of tracking students across multiple institutions, according to a recently completed survey by the State Higher Education Executive Officers. At the same time, an organization called the National Student Loan Clearinghouse is tapping financial aid records on a national basis to allow former students to be followed almost anywhere.

Virtually all states now have in place or are constructing unit-record enrollment databases capable of tracking students across multiple institutions.

Assessment Update
January-February 1998
Volume 10, Number 1

BEST COPY AVAILABLE

As signaled by a meeting in August 1997 in Washington, D.C., convened by Mathematica Corporation, with representation drawn from a number of states and federal agencies (including the Departments of Education, Labor, and Commerce), the character and speed of these developments are beginning to raise major issues of public policy. On the one hand, virtually all these record-linking initiatives have occurred independently, with resulting inconsistencies in both their data content and in the ways performance indicators like placement or graduation rates are constructed. Most states, for instance, don't yet include detailed occupational data in their unemployment insurance records—substantially limiting the ability to create indices of "placement related to field of study." One purpose of the gathering in Washington, D.C., therefore, was to explore proposed common efforts to expand the coverage of these databases, as well as to extend technical assistance efforts designed to bring new state and federal users into the fold. On the other hand, the increasing use of such databases also raises substantial concerns about privacy, which in turn are revealing interesting contradictions in reigning conservative political ideologies. Groups like the American Eagle Forum, for instance, have aggressively fought the use of individual record information on libertarian grounds and have actually succeeded in reversing student-tracking efforts in Oregon through court order. But conservative politicos also increasingly emphasize the need for more specific performance reporting at both the state and federal levels and are highly sympathetic to constructing such measures in the cheapest ways possible. One result is a lot of misinformation (as well as some deliberate disinformation) about what one can and cannot do in this arena. Guidelines provided by the Federal Information Rights and Privacy Act and a bill on computer matching recently passed by Congress, for instance, strictly govern the use of the Social Security number (SSN) to track individuals for the purposes of high-stakes intervention (for example, to identify loan defaulters or those behind in their child-support payments). But they are far less constraining on the use of SSNs as a tool to link databases for aggregate research purposes. An interesting grey area here is the use of aggregate indicators constructed from such records in making high-stakes judgments of *institutional* performance. In Florida, for instance, a lot of the pressure to link databases in the first place is being driven by policy interest in performance funding.

Applications of technology to instructional delivery are forcing greater attention to the shape of national data collection.

In response to all of these issues, leading states are again attempting to apply technology. Both Florida and Texas have created sophisticated computer environments that (so far) can effectively "firewall" access to individual record information but that still allow a range of client groups to conduct aggregate-level studies that depend on record linking. The political battles surrounding these issues, however, are just getting started.

Applications of technology to instructional delivery—particularly in distance education—are at the same time forcing greater attention to the shape of national data collection. The focus here, though, is the growing inability of established national data definitions and data-collection approaches like the federal Integrated Postsecondary Education Data System (IPEDS) to reflect the complexities of enrollment, progression, and attainment in a virtual instructional environment. As but one extreme example, under

Assessment Update
January-February 1998
Volume 10, Number 1

current IPEDS definitions the nascent Western Governors University will report no enrollments but substantial numbers of graduates. Such cases, to say the least, render absurd compliance with the established graduation rate provisions of "Student Right-to-Know."

Yet another August gathering in Washington, this one under the auspices of the National Postsecondary Education Data Cooperative, brought together research and government representatives to explore this issue more explicitly. Although somewhat futuristic, two prominent themes of this discussion should be of substantial interest to assessment practitioners. One is the growing need to reconceptualize the structure of federal databases from a largely cross-sectional (snapshot) architecture that uses *institutions* as the primary unit of analysis, to a longitudinal configuration that uses *individual students* as the principal unit of analysis. This need, in turn, can be attributed to the increasing ability of students to access multiple providers simultaneously, as well as to increasingly episodic patterns of enrollment. Using record links among existing national databases described earlier to carry out this kind of follow-up is, of course, a major consideration. A second prominent theme is the growing obsolescence of time-based measures of student progress and attainment (for example, the credit or clock hour), which currently constitute the cornerstone of federal enrollment definitions. And though no one can claim to know how a transition might be managed, the principal conceptual alternative to such measures are *competency-based* benchmarks of some kind.

Changing technical capacities in themselves will be highly influential in shaping future accountability requirements.

These developments support the view that changing technical capacities in themselves will be highly influential in shaping future accountability requirements. At the same time, they suggest no loss of interest among those in authority in creating such requirements. But what may actually come to the surface in the wake of the likely tempests in this brave new world, only time will reveal.

Implementing Performance Funding in Washington State: Some New Takes on an Old Problem

The emerging "second wave" of performance funding initiatives is a good deal more sophisticated than the last.

Half a decade has gone by since states began joining Tennessee in using performance measures to allocate some portion of their public higher education dollars. Kentucky, Arkansas, and Missouri were among the first to do so. Now about ten states run formal performance funding programs, and a dozen more are seriously considering their adoption. Although for many this trend is worrisome, the emerging "second wave" of performance funding initiatives is a good deal more sophisticated than the last. Among the most interesting is under way in the state of Washington.

The political appeal of pay for performance remains undiminished. Indeed, enthusiasm for it has thus far spawned three attempts to allocate *all* public dollars to colleges and universities on this basis—two patently unsuccessful (in Texas and Florida) and one still under way (in South Carolina). But less extreme approaches have also proven difficult to implement. Once in place, moreover, it is hard to make them stick. At least one early adopter (Arkansas) is now out of business; as of this writing, it will likely be joined by another (Kentucky). For the most part, these difficulties are a product of sheer political instability. Implementing a successful performance funding policy—as Tennessee's fifteen-year experience graphically demonstrates—is a long-term business and is rarely gotten right the first time. But in many states these days, policymakers either cannot agree on the common goals needed to drive a performance funding program in the first place or cannot maintain consensus long enough to resist fatal meddling after a program is established.

Recent experience in Washington nicely illustrates many of the policy dilemmas involved. The state's approach also contains both innovative and problematic features that are present in other "second wave" performance funding states. The story begins with the Washington legislature's decision in 1997 to include a performance component for higher education in the state's 1996-1998 biennial budget (ESHB 2259, Sections 601 and 619). Governor Gary Locke was an active proponent of this initiative. Earlier, as a state legislator, Locke had been a visible supporter of the state's approach

Assessment Update
May-June 1998
Volume 10, Number 3

BEST COPY AVAILABLE

to assessment. What happened in assessment, in fact, helps explain what happened in performance funding—largely because of the policy tone that the earlier initiative established.

In 1989, Washington institutions successfully resisted a mandate by the Higher Education Coordinating Board (HECB) for a commonly administered sophomore test, substituting instead a policy based on local, institutionally developed measures. More significant, the legislature generously supported this more decentralized initiative through appropriations to individual colleges and universities. As a result, Washington is still one of the few places where mandated assessment activities are supported by state funds. Issues raised by assessment—institutional autonomy, the use of categorical funding to induce institutional action, and the policy roles of intermediate bodies like the HECB and the State Board for Community and Technical Colleges (SBCTC)—would all be revisited in acting out performance funding.

One important difference between Washington's approach and those of other states is that it was decentralized from the outset. In fact, the state has not one but two performance funding initiatives. For the six four-year institutions, the set-aside pot was designated to be 2 percent of the noninstructional budget, but because each institution is independently governed, incentive dollars were separately appropriated for each. To allocate these funds, five performance measures were established by law: (1) undergraduate student retention rates, (2) five-year graduation rates, (3) performance on a "graduation efficiency index" based on the ratio of actual credits completed and the number officially required for program completion, (4) a faculty productivity measure uniquely defined by each institution, and (5) an additional uniquely determined measure reflecting "mission attainment" for each institution. The HECB was charged with establishing rules for determining how much each institution could claim of "its" performance dollars, but was not itself given the money.

One important difference between Washington's approach and those of other states is that it was decentralized from the outset.

As a recognized governing body, though, the SBCTC was given both the money and the authority to allocate it directly to the state's thirty-two two-year institutions. From the outset, the SBCTC was thus able to adopt a more flexible approach, taking into account appropriate differences between institutions. In this case, the performance pot was established at 1 percent of total appropriations for the system as a whole. To allocate these dollars, explicit performance targets were set for four statewide outcomes: (1) the hourly wage of job-training program graduates, (2) the transfer rate of academic-track students to four-year institutions, (3) completion rates in core courses such as English and math, and (4) performance on the same "graduation efficiency index" used by the four-year institutions.

Both initiatives, though, involved taking money off the top-a strong break with established Washington tradition, as well as with the experiences of most other states using performance funding. One way or another, most of the latter have managed to create new money to support their ventures (or at least to create the appearance of doing so by establishing a visibly separate source of funds). Washington's experience confirms that the initial tone set by such moves is important. A more creative initial funding approach might have avoided the strong opposition to the measure by campus leaders, especially in the four-year sector, where the signal sent by the legislature implied restricting money that had already been appropriated.

Despite this drawback, both the SBCTC and the HECB have in part been able to adopt approaches that buffer the "hard linkages" between statistical performance levels and dollar amounts that are typical of other performance funding schemes. Hard linkages have always presented two fundamental problems. One is the lag effect inherent in almost all performance measures. Even if an institution takes immediate action to address a particular outcome (for example, a graduation or retention rate), it may be years before these actions pay off. Meanwhile, the institution continues to be penalized in an outcome-based performance scheme.

Even if an institution takes immediate action to address a particular outcome, it may be years before these actions pay off.

To address this difficulty, both the HECB and the SBCTC in Washington awarded performance funds for the first year of the biennium solely on the basis of institutional planning efforts aimed explicitly at improving performance on the statewide outcome measures. Going further, the SBCTC approach also required each two-year college to adopt a set of additional action or process measures, chosen from a statewide menu, that were expected to contribute to collective performance on these outcomes. Institutional fulfillment of locally established targets for these action measures would then be used by the board to help allocate performance dollars in the second year of the biennium.

The second historic problem with the use of hard linkages in performance funding is how to take into account factors outside an institution's control. Graduation and retention rates, for instance, are determined far more by incoming student ability levels than by anything that an institution can do. Institutions also may simply be unable to benefit from particular performance "opportunities" because of where they are and what they do—a difficulty apparent, for example, in Missouri's Funding for Results (FFR) program that rewards institutions explicitly for producing minority graduates. Two useful ways to address this problem are to focus principally on institutional improvement and to negotiate with institutions to establish individualized (and appropriately different) expectations. Both have been partially adopted by the HECB and the SBCTC. In the latter case, the board works directly with institutions to set appropriate targets for attaining statewide performance values, taking individual differences in context and capacity into account. In the four-year case, this approach is less feasible because *all* institutions are expected to attain the same statewide performance targets by 2005. There is some flexibility even here, however, because different performance increments have been negotiated for each institution for each intervening year.

A final promising feature of both Washington initiatives—pioneered by Missouri's FFR program several years ago—is their use of additional institution-specific measures that can be tailored to particular missions and contexts. As noted, the four-year program includes two institutionally defined measures in the group of five established in law. The SBCTC, in

Assessment Update
May-June 1998
Volume 10, Number 3

Statewide Testing: The Sequel

Lately statewide standardized testing has been staging a comeback.

When states began jumping on the mandated assessment bandwagon in the mid-to-late eighties, the first option proposed was always standardized testing. The reasons were many but always included legislative familiarity with standardized assessments in K-12, ease of acquisition, and the immediate public credibility that accompanies any instrument presumed to be valid and reliable. Counter-arguments from higher education institutions in this period were vehement and predictable. More important, they were largely successful in diverting most mandates toward institution-centered approaches. Instead of forcing all institutions to adopt a single instrument through which their relative standings could be compared, colleges and universities were in most cases given the freedom to select their own assessment methods, as long as they did so responsibly, and publicly reported the results they obtained. Prominent milestones during this time of troubles were Washington's large-scale study of existing standardized tests in general education, which rejected all available alternatives as unsuitable; New Jersey's innovative College Outcomes Evaluation Program, which replaced multiple-choice tests with authentic (but expensive) task-based assessments; and South Dakota's abandonment of statewide general education testing as politically ineffective after three years. By the early nineties—with more than two-thirds of the states in the assessment business—only a handful maintained a standardized testing program; among stalwarts, moreover, were states such as Florida, Tennessee, and Georgia, which launched programs well before assessment became an accountability fashion, and undertook them for quite different policy reasons.

Lately, however, statewide standardized testing has been staging a comeback. In 1996, for instance, South Dakota again reversed its position, requiring all students to take the ACT-CAAP test. Similar and somewhat older programs in Arkansas and Wisconsin also use the CAAP. Most recently and visibly, this spring the State University of New York (SUNY) system and the Utah Board of Regents declared strong interests in establishing standardized testing programs for accountability purposes—in the former case, even issuing a widely circulated request for proposal (RFP) to assessment

Assessment Update
September-October 1998
Volume 10, Number 5

BEST COPY AVAILABLE

turn, requires each institution to identify an individualized set of process measures for which to be held accountable in the biennium's second year. The board may also allow colleges to set their own priorities regarding which of the four legislative outcomes measures they want to emphasize (and be rewarded for) in their local improvement plans.

On balance, the notion of establishing a few broad performance goals and allowing established governing or coordinating bodies the flexibility to allocate resources to institutions at different levels and in different ways to attain these targets as a *system*, seems a promising innovation for performance funding. So far, Washington's approach has come closest to testing this notion in practice. Colorado may adopt a similar approach next year. And although South Carolina's far more ambitious performance funding adventure will continue to grab most of the headlines, we may eventually learn just as much from the opposite corner of the country.

Establishing a few broad performance goals and allowing established governing or coordinating bodies the flexibility to allocate resources to institutions at different levels and in different ways to attain these targets as a system, *seems a promising innovation for performance funding.*

providers before a decision was made. Four of five of these initiatives emanate from governing boards rather than from coordinating boards (the source of the Arkansas initiative) or legislatures. All five make clear their intent to judge institutions comparatively; unlike Texas or Florida, for instance, no states, except South Dakota, will use test results to affect the fates of individual students. Finally, all five states make no bones about using off-the-shelf instruments; the SUNY RFP, for instance, requests that providers propose programs that can be fielded quickly by explicitly specifying the use of existing tests.

What lies behind this growing level of interest and activity? One part of the explanation is that institutional accountability measures in general—chiefly in the form of performance indicators—have grown sharply in popularity across the states. At least twenty states now report such statistics annually for all public institutions, and about eight reward high-scoring institutions directly with performance or incentive funds. With the exception of Tennessee, however, few states have actually used the results of outcomes testing to allocate funds. Some of the reasons for this anomaly were captured succinctly in national surveys conducted by NCHEMS and the University of Michigan in 1966-97: although most state officials wanted outcomes testing as part of their performance-measure systems, at the time of these surveys few were ready to bear the fiscal and political price of trying to implement a large-scale testing program. With state budgets now bouncing back, accountability as salient as ever, and a new breed of substantially more aggressive incumbents sitting on state higher education boards, these restraining conditions are rapidly evaporating.

The new breed of politicians involved in these decisions is also a factor in itself. On the one hand, term limits and a radically changed political balance in most statehouses mean that few sitting legislators or board members remember the relatively sophisticated debates about the merits and demerits of standardized testing that took place a decade ago. Unacquainted with alternatives, most such legislators naively see tests as promising and uncomplicated policy tools—just as their predecessors initially did. More ominously, many governing boards are now populated by a new cast of characters—actively political and quite explicit in their desire to dictate what is taught in public colleges and universities. SUNY's interest in standardized testing, for instance, is part of a wider pattern of proactive involvement in curricular matters designed to reverse what many board members feel has been a dangerous erosion of academic standards. Holding campuses accountable for educational results through testing is seen as one of many steps—including required core courses and incentive budgeting approaches—that will both increase system productivity and ensure that students are being taught what they should know.

One somewhat surprising development is that this new wave of testing proposals has so far encountered neither righteous nor reasoned resistance from the academy.

These reasons for interest in standardized testing raise some intriguing issues, however. One somewhat surprising development is that this new wave of testing proposals has so far encountered neither righteous nor reasoned resistance from the academy. Both reactions, of course, were characteristic of prior attempts to enact such measures. At SUNY, to be sure, faculty union representatives raised the usual objections to testing before the board in the spring, but grassroots opposition seemed lackluster in comparison to what happened in New Jersey and Washington years ago. In Utah, few at the universities (at least as of this writing) appear either threatened or excited by the initiative. Part of the reason in both cases may be a growing ho-hum attitude toward accountability in general that's the product of more than fifteen years of sustained government pressure: many academic leaders may have come to believe that, as with taxes and pollution, nothing can be done except to live with it.

At the same time, the essentially conservative educational and political philosophy that undergirds current initiatives contains real dilemmas about how a testing program ought to be organized. On the one hand, the creation of national K-12 standards, backed by assessment, is exactly the agenda now being advanced by the Clinton White House—and being vehemently opposed by conservative leaders on the grounds of state and local rights. Indicative here is the fact that contemporary Republican-led fiscal proposals in New York that would create clear winners and losers among SUNY institutions are already being actively opposed by equally conservative legislators whose own campuses might be adversely affected. More important, the academic content being advocated by the New Right as the foundation of reinvigorated standards is not what available standardized instruments test. Rooted in the assessment demands of the eighties, examinations like the COMP, the CAAP, the C-BASE, and the Academic Profile focus far more on cross-cutting academic skills such as communication and critical thinking than on knowing the traditional canons of literature, science, and Western civilization." Building new tests will cost money, however—exactly what efficiency-minded board members want to avoid.

Statewide testing proposals may prove immensely popular among policymakers yet be impossible to implement because sufficient long-term consensus cannot be forged to make them work.

Meanwhile, the realities of implementing large-scale testing programs in college and university settings remain as intractable as ever. Both Utah and SUNY—like Arkansas and Wisconsin—seem unwilling so far to make test performance a condition for student advancement. Problems of motivation under such conditions have been apparent from the beginning and will probably continue to be encountered in such programs. Similarly, both states' proposals recommend testing samples of students rather than full populations. Again, experience elsewhere suggests that once a state is committed to such a program, all affected students will have to be tested, if only on grounds of equity and logistical convenience. Again, these practical problems lead to higher costs, and sustained attention and unbroken political will will be required to overcome them.

Neither of these intangibles is in great supply in the states these days. As is also the case with complex performance funding schemes, statewide testing proposals may prove immensely popular among policymakers yet be impossible to implement because sufficient long-term consensus cannot be forged to make them work. Equally relevant is the fact that the conservative standards revolution that lies behind recent testing proposals is really directed toward what is taught and how, not toward how much is learned. If sufficient political capital is mustered to sustain a change effort, it will likely be directed more

Assessment Update
September-October 1998
Volume 10, Number 5

toward the heart of the curriculum than toward launching a complex and costly testing program.

Whatever the outcome, the latest round of statewide testing proposals is disturbing, if only because we have been through all this before. Like most summer remakes, though, we can probably count on statewide testing's sequel this year to be much like its fellows in the movie industry: cheaper, shorter, and a good deal less subtle and sophisticated than the original.

The latest round of statewide testing proposals is disturbing.

Assessment Update
September-October 1998
Volume 10, Number 5

Community College Strategies

Jeffrey A. Seybert

About the Author

Jeffrey A. Seybert is currently Director of Research, Evaluation, and Instructional Development at Johnson County Community College in Overland Park, Kansas (in the Kansas City, Missouri metropolitan area). He has been in the Office of Institutional Research at JCCC for more than seventeen years. Prior to his appointment at JCCC, Jeff served as assistant professor of psychology at the University of Missouri-Kansas City. Jeff holds a bachelor's degree in psychology from Cal-State Long Beach, master's degrees in public administration from the University of Missouri-Kansas City and experimental psychology from the University of Oklahoma, and a doctorate in experimental psychology from the University of Oklahoma. He is past president of the National Council for Research and Planning and the MidAmerica Association for Institutional Research, and has served on the Board of Directors of the Association for Institutional Research (AIR). Jeff serves as consulting editor for the *Journal of Applied Research in the Community College* and *The AIR Professional File*, and consulting editor and contributing columnist for *Assessment Update*. He has written numerous articles and book chapters and served as consultant to approximately 85 colleges, non-profit and other organizations in the U. S., Great Britain, Australia, and the Caribbean in the areas of assessment of student outcomes and institutional effectiveness, program evaluation, planning, and institutional research.

Contents

Page

Community College Strategies 245

Assessment Issues, Methods and Policies 249
Models for Assessing Institutional Effectiveness 251
How to Initiate an Assessment Program 253
Assessment of the Transfer Function 255
Assessment of Career and Occupational Education 258
Assessment of General Education 261
Assessing Student Learning 264
Assessment in Washington State Community and Technical Colleges 267
Assessment of Noncredit Continuing Education and Community Service Programs and Courses 271
Community College System-Level Approaches to Core Indicators of Effectiveness 274
Virginia Community College System 278
Assessment at Two-Year Institutions 280
Developing a Statewide Core Indicator Project for Community Colleges 282
A New Paradigm for Evaluating Transfer Success 285
The Wisconsin Technical College System Institutional Effectiveness Model . . . 288

JEFFREY A. SEYBERT

Community College Strategies

Looking back on the *Community College Strategies* columns since their inception in the July-August 1992 issue of *Assessment Update* provided me with an interesting opportunity to review the content of those essays, attempt to organize them in some way that made sense, think about additional topics we should or might need to address, and in general, reflect on what direction the column might take in the future.

As I looked through the essays, they seemed to almost naturally organize themselves in two groups. First, taken together, those I'd written could almost make up a brief, cursory "primer" of community college assessment, with short "chapters" on models for assessing institutional effectiveness; how to initiate an assessment effort; assessment of the transfer, career/occupational, and continuing/community education functions; and assessment of student learning in general, and general education specifically. The information in those essays still strikes me as sound, if somewhat abbreviated, coverage of those topics.

The essays almost make up a brief, cursory "primer" of community college assessment.

On the other hand, the columns written by guest authors seem to focus either on state and national perspectives on assessment in community colleges, or on proposed innovations in assessment practice. For example, the essays by Bill Moore (6:6), Debra Banks (8:2), Rick Voorhees (9:2), and Don Bressler and Deborah Mahaffey (10:4) reported on a variety of assessment issues and practices in Washington, Virginia, Colorado, and Wisconsin, respectively. On a national level, Mike Smith and Jama Bradley (8:4) described the two-year college results of a national survey of assessment practices conducted by the Clearinghouse for Higher Education Assessment Instruments at the University of Tennessee, Knoxville while Loretta Seppanen (7:6) discussed results of a survey of state community college offices conducted by the Washington State Board for Community and Technical Colleges to determine the extent to which state-level community college organizations are involved in analyzing and publishing core indicators of effectiveness for the public two-year sector. Finally, a piece by Craig Clagett (7:4) proposed a new outcomes typology for community colleges, and an essay by Mike Quanty and his

BEST COPY AVAILABLE

colleagues (10:2) described the development and pilot test of a "course-based" model to assess transfer success. In my opinion these essays provide a fascinating look at the wide variety of assessment activities taking place in and for community colleges.

Reviewing these essays gave me the opportunity to reflect on the "state of the art," assessment-wise, when the *Community College Strategies* (CCS) column was initiated, what I and other authors have written about assessment since the column originated, and how assessment practices and policies have changed for community colleges in those years.

In some instances, we were right on target. As I noted in the very first CCS column in 4:4, assessment is different for two-year colleges than for four-year institutions, and that remains the case today; although, I've come to realize this issue is more complex than I understood at the time. Due to the somewhat transient nature of many of our students, the diversity of their educational goals, and their penchant for transferring before we know they're going to do so, we still have considerable difficulty in identifying, "capturing," and motivating them to perform well on assessments of learning. For a variety of reasons, we cannot impose participation in the assessment process as a graduation requirement (as do many of our four-year college and university counterparts), so standardized testing, for example, becomes quite problematic for us; although it remains the preferred method for assessing student learning in many community colleges. In fact, together, standardized and locally created testing instruments are the methodologies of choice to assess student learning in a majority of community colleges, although there is a growing trend toward authentic or performance-based assessments.

Assessment is different for two-year colleges than for four-year institutions.

On the other hand, my experience since I wrote that column has shown me that our colleagues in four-year institutions have quite a bit more difficulty involving faculty in the assessment process than do we. Given that the primary role of community college faculty is centered on teaching and learning, it's not surprising that they more readily participate in the assessment process than do their four-year brethren, who also are responsible for research, publishing, and other scholarly activities.

In 5:4 I noted some of the difficulties in attempting to assess the effectiveness of the transfer function. Specifically, obtaining identifiable individual student record information from four-year institutions is a critical piece of data if we are to assess how well we prepare students for specific transfer majors. For the most part, the intervening years have not brought a solution to these problems. In fact, in some cases, such data exchange has been expressly prohibited (e.g., in Kansas) by university legal counsel. Thus, alternative methodologies for assessment of the transfer function, such as the course-based model of transfer success suggested by Mike Quanty and his colleagues in 10:2, are most welcome.

Assessment of the career/occupational function, as I discussed in 5:6, is still a very important component of two-year colleges' comprehensive programs to assess institutional effectiveness, as evidenced by the columns in 9:2 by Rick Voorhees and in 10:4 by Don Bressler and Deborah Mahaffey, dealing with statewide assessment systems in the Colorado community colleges and Wisconsin technical colleges, respectively.

Assessment of general education, as I discussed in 6:2, is still very high on the radar screen for community colleges. General education outcomes, and various alternative ways to measure them, are topics of major concern as colleges continue to design, refine, and implement their programs to assess institutional effectiveness and student learning. While standardized and locally produced tests continue to predominate as assessment methodologies of choice for general education outcomes, alternatives to testing, such as individual and institutional portfolios (including the model described in the article in 9:4 by my colleague Kathy O'Hara and me), capstone courses, competence checklists, and even analysis of course syllabi (as reported by Trudy Bers and her colleagues in 8:6) are becoming increasingly popular.

While most two-year colleges have made important strides in the assessment areas noted above, as well as in assessment in the major, there are others in which much work remains to be done. For example, notwithstanding the discussion of assessment of continuing education/noncredit/community service programs in 7:2, we still have a long way to go in terms of assessing effectiveness of these types of offerings. In addition, there are other important components of most community colleges' missions to which scant attention has been paid, assessment-wise. These include economic development, diversity and multiculturalism, noncognitive/affective student growth and development, and culture and ethics.

A major challenge is and will continue to be sustaining faculty interest, enthusiasm, and involvement in assessment.

There are also several assessment issues concerning two-year colleges that have dramatically increased in importance since the column was initiated. The most notable of these is the emergence of statewide performance or core indicator systems, as evidenced by the column in 7:6 by Loretta Seppanen describing a national survey of community college state system offices concerning core indicators, Rick Voorhees' column in 9:2 dealing with the development of a set of statewide core indicators for community colleges in Colorado, and Don Bressler and Deborah Mahaffey's column in 10:4 dealing with the design and implementation of a statewide system of core indicators for the Wisconsin Technical Colleges.

Another theme that has become more clear since the inception of the CCS column is that the impact of assessment on curricula, teaching, and learning will take considerable time to make an impact. Many faculty and assessment practitioners hoped and/or expected (now, it seems, a little naively) that we would see the results of assessment in a relatively short timeframe. It has become clear, however, that the systemic organizational change we hope assessment will produce will take considerably longer to materialize than was initially thought. Thus, a major challenge is and will continue to be sustaining faculty interest, enthusiasm, and involvement in assessment. Finally, assessment of distance learning programs will become a topic of increasing importance as those programs continue to proliferate.

In addition, we will continue to wrestle with issues such as faculty resistance; incorporation of assessment (and assessment results) into ongoing institutional processes like planning, program review, and curriculum development and modification; overall institutional resistance to change; involving adjunct faculty; and finding the resources necessary to carry out assessment programs.

Finally, as I reviewed all these essays, it struck me that one of the major strengths of a regular column like *Community College Strategies* is the opportunity it gives all of us who are working with or interested in assessment in two-year colleges to benefit from the experience and expertise of our very talented colleagues. Tremendous progress has been made toward our goal of accurately measuring our institutions' effectiveness and what our students learn, and then using results of those assessments to improve teaching, learning, and delivery of services to students. I would hope that the column will continue to reflect that progress. One way to insure that, of course, is to continue to receive timely information from those on the front lines of the assessment effort; I trust we will be able to do that. In addition, I think future columns should discuss assessment of important components of the community college mission that we have not yet addressed, like remedial/developmental education, economic development, access, diversity, distance education, and others. I would welcome essays on any of these topics.

Tremendous progress has been made toward our goal of accurately measuring our institutions' effectiveness and what our students learn, and then using results of those assessments to improve teaching, learning, and delivery of services to students.

On reflection, it has been a privilege to be involved with *Assessment Update*; to have had the opportunity to share some of my experiences with those interested in assessment in two-year colleges, and to have worked with the group of very talented higher education professionals who are currently involved with assessment in community colleges. The *Community College Strategies* column has benefited from their wide-ranging expertise. And, finally, I would like to invite readers to submit their ideas for future columns. We can learn much from the experiences of our colleagues and I would very much appreciate the opportunity to showcase those experiences in the *Community College Strategies* column.

JEFFREY A. SEYBERT

Assessment Issues, Methods and Policies

This is the first in what is planned as an every-other-issue column dedicated to assessment in community colleges. The column will focus on issues, methods, and policies that are specific to two-year colleges, although they may have broader appeal and interest as well. The column will not duplicate Peter Gray's, which is designed to highlight assessment efforts and programs on specific campuses, including community colleges. Although activities that work particularly well on a given campus may be described here, the emphasis will be to describe those processes rather than to profile an entire institutional effort.

I invite contributions from anyone interested in assessment in two-year colleges. These contributions may include examples, comments, suggestions, and even guest columns about assessment issues. In fact, I hope that some items in the column will be provocative and elicit discussion, argument, and rebuttal. Feel free to disagree with what you read, and let me hear your opinions.

Assessment should relate to institutional effectiveness. In this first column I want to address briefly three issues. First, I believe that assessment in two-year colleges should be a broad approach to evaluation of overall institutional effectiveness. This is not to diminish the importance of assessing student learning outcomes. Rather, outcomes assessment is one of a number of crucial components of a larger effort to examine all facets of the two-year college and its diverse goals. Assessment of learning outcomes is probably the most difficult, albeit one of the most important, of these components. Thus, although community college faculty and staff must significantly increase their efforts in outcomes assessment, they cannot let this increased emphasis supplant existing and planned efforts to assess other facets of their institution.

Assessment is different for two-year colleges. Second, the differences in assessment in two- and four-year institutions need to be recognized. Demonstrating institutional effectiveness presents a special set of problems for community colleges, which typically have a much broader instructional mission than do four-year colleges and universities. In addition to traditional freshman- and

Assessment in two-year colleges should be a broad approach to evaluation of overall institutional effectiveness.

Assessment Update
July-August 1992
Volume 4, Number 4

BEST COPY AVAILABLE

sophomore-level coursework, community colleges provide career training, occupational retraining, remedial and developmental coursework, community and continuing education programs, contract training for business and industry, courses for special populations, and a variety of other educational offerings. Also, community college students often differ dramatically from traditional college students. They are more diverse in terms of age, background, employment status, preparation, and educational objective than their four-year college and university counterparts. Thus, some measures of institutional effectiveness commonly used by four-year colleges and universities, such as number of graduates or proportion of graduates to students admitted, are not usually applicable to community colleges. For these reasons, those of us charged with the responsibility of assessing the effectiveness of two-year colleges face an especially difficult challenge.

Assessment is not . . . Third, it is important to identify several things that assessment isn't. As I define it, assessment is not individual performance evaluation. Many faculty, in particular, seem to fear that the results of assessment will somehow be tied to their individual performance and, further, to salary, promotion, and tenure decisions. It is critical to define very clear and explicit boundaries between assessment and performance evaluation and to observe these boundaries strictly. Faculty buy-in and participation are crucial to the success of any assessment effort. The quickest and surest way to guarantee failure is to permit the faculty to perceive that they will be evaluated based on assessment results. Administrators must make a firm commitment to the absolute dissociation of assessment and performance evaluation.

A clear understanding of the purposes for and boundary conditions of assessment will allay the fears that many faculty and administrators harbor regarding implementation of assessment programs.

Assessment should not generally be considered the evaluation of outcomes at the course level. Assessment of institutional effectiveness clearly involves assessment of overall student learning outcomes, both in terms of general education and at the departmental or discipline level. However, determination of course-level outcomes is the sole purview of the individual faculty member and should take place as part of course assignments, exercises, papers, and examinations. Individual faculty members may, of course, choose to undertake classroom assessment projects on their own. However, institution-wide assessment of student learning outcomes should take place at the curricular/departmental/discipline level or above.

Assessment also is not, if appropriately designed and conducted, an infringement on academic freedom. The principle of academic freedom was established to protect faculty who hold or express political, social, and other views outside accepted or commonly held norms. It should be clear that assessment has no impact on these behaviors. In addition, assessment is not designed to dictate to individual faculty members either what they teach or which instructional methodology they use. Finally, the earlier arguments for explicit dissociation between assessment and performance evaluation and between assessment and evaluation of individual course outcomes reinforce the position that assessment does not violate academic freedom.

A clear understanding of the purposes for and boundary conditions of assessment will allay the fears that many faculty and administrators harbor regarding implementation of assessment programs.

Assessment Update
July-August 1992
Volume 4, Number 4

JEFFREY A. SEYBERT

Models for Assessing Institutional Effectiveness

Planning and implementing a program to assess institutional effectiveness and student outcomes requires a model to provide an underlying conceptual framework or context; to organize, focus, and guide assessment; and to ensure that the assessment effort is comprehensive and implemented in all operating units. Four models have been suggested to guide assessment in community colleges.

Four models have been suggested to guide assessment in community colleges.

Probably the first such model was proposed in the early 1980s by members of the National Association of Community and Technical Colleges with assistance from staff at the National Center for Research in Vocational Education, then at Ohio State University. Thirty-nine institutional effectiveness indicators were proposed in seven specific areas, including: access and equity, employment preparation and placement, college/university transfer, non-degree college services, economic development, community involvement, and cultural and cross-cultural development.

This model has been adapted for use at several community colleges, notably Midlands Technical College in South Carolina, Utah Valley Community College, and some California community colleges.

A second model was developed by Jim Nichols at the University of Mississippi with input regarding implementation in community colleges from Harriet Calhoun at Jefferson State Community College in Alabama. The Nichols model has four major elements:

- Establishment of an "expanded" statement of institutional purpose
- Identification of intended outcomes for all functions of the institution (instruction, student services, administrative services, support services)
- Assessment of the extent to which these outcomes are achieved
- Adjustments in the institution's purpose, intended outcomes, or activities based on the assessment results.

Assessment Update
November-December 1992
Volume 4, Number 6

Because of the scope, time intensity, and sequential nature of the component of Nichols' model, he recommends implementation over at least a four-year period. The model has been adopted at numerous institutions, including a modified version at Jefferson State.

A third model has been developed by a task force of the League for Innovation in the Community College and published as an edited monograph, *Assessing Institutional Effectiveness in Community Colleges*. The monograph focuses on (but the model is not limited to) five key missions of most community colleges: transfer, career, basic skills and developmental education, continuing education and community service, and access (labeled the "superordinate" mission for most community colleges). The model involves identifying the major components of a college's mission, articulating key or seminal questions (outcomes) for each component, and identifying appropriate measures to answer each question or assess each outcome. This model has been adopted, at least in part, in a number of institutions, including St. Louis Community College.

It is up to each institution to pick the model most appropriate for its situation and personnel and then tailor that model to meet its unique needs and mission.

I have proposed a fourth model, the Effectiveness Assessment Matrix (EAM), a two-dimensional, four-celled matrix with variable or outcome (student versus institutional) and audience for results (internal versus external) serving as the dimensions. The matrix thus yields four categories of effectiveness measures: internally directed student measures (such as course grades, grade-point averages, satisfaction of individual educational objective), externally directed student measures (student performance on professional licensure exams, transfer student success in the four-year college or university), internally directed institutional measures (results of systematic program reviews, analyses of course grades and retention/attrition rates), and externally directed institutional measures (results of employer evaluations of student preparation and other external studies). The model can be applied to all the components of a college's mission to yield a comprehensive assessment of each component's effectiveness. This approach has been adopted, all or in part, at several institutions, including Johnson County Community College.

An important feature shared by all these models is the emphasis placed on the primacy of the college mission and mission statement as a foundation on which to base a program to assess institutional effectiveness. Each model also affords considerable flexibility in the design and implementation of an individual college's assessment effort. Using any of these models thus permits customization of such an effort, allowing a college to design its assessment program specifically tailored to its needs, particular circumstances, and unique mission. Finally, I want to emphasize that I am not making any claims about the relative efficacy of these models; each will do the job well. It is up to each institution to pick the model most appropriate for its situation and personnel and then tailor that model to meet its unique needs and mission.

Assessment Update
November-December 1992
Volume 4, Number 6

JEFFREY A. SEYBERT

How to Initiate an Assessment Program

As I travel around the country talking to faculty and administrators at two-year colleges about assessment of institutional effectiveness and student learning outcomes, the single question that I am asked most often is "How do we start—what should we do to begin to implement an assessment process?" Thus, I present here a brief list of steps that I believe an institution may want to take to initiate an assessment program.

What should we do to begin to implement an assessment process?

1. *Create an institutional effectiveness assessment task force.* This task force should be broadly representative of college faculty and administrators and should include (but not be limited to) a senior administrator (president, vice president, or dean of instruction) firmly committed to assessment, institutional research personnel, faculty, and student affairs staff. Committed faculty and upper-level administrative involvement are absolutely critical: I am firmly convinced that assessment cannot be successful without them. In addition, I think that this should be a permanent task force whose job is to guide, evaluate, and modify the assessment effort as it evolves.

2. *Examine the college mission statement.* Assessment is essentially an examination of the degree to which the institution is in fact adhering in practice to the principles of its mission statement. Since the college mission statement is the standard against which institutional effectiveness will be measured, it is critical that the mission statement be up to date and fully reflect what the college is supposed to be doing. This examination may dictate a revision of the mission statement, depending on its currency.

3. *Design an institutional plan or model to guide the assessment program.* This plan may (or may not) be based on one of the four formal models that have been proposed as conceptual frameworks for assessment (see the *Community College Strategies* column in *Assessment Update*, 1992, 4(4), 13). In any case, it is helpful to have a plan or model to guide the assessment effort so as to ensure that it addresses all major facets of the institution's activities and places assessment in some overall context.

Assessment Update
March-April 1993
Volume 5, Number 2

4. *Determine and prioritize the specific assessments to be undertaken.* The priorities will depend on the immediacy of external mandates, internal needs for assessment data, and the like.

5. *Inventory existing data collection efforts.* Even in institutions with no formal institutional research or assessment office, numerous data collection efforts are underway. Many of these can be incorporated into the assessment process.

6. *Determine what additional data collection procedures need to be implemented to inform the assessment priorities identified in step 4.* This step may require the implementation of additional data collection efforts. Implementation of an overall assessment effort will not generally occur without some new costs.

7. *Start at the top of the priority list.* Start small, with selected pilot projects in areas where faculty and staff are excited and ready to go.

8. *Be flexible, adaptive, and prepared to change.* There will always be problems; mandates, circumstances, personnel, and priorities can always change. Assessment must be viewed as dynamic and evolutionary and not as a static process that is set in concrete once it is fully implemented.

9. *Be prepared and willing to publicly share results of assessments.* Experience shows us that the results of these assessment procedures are almost always positive. They must provide marvelous public relations opportunities and also help reinforce the commitment of faculty and staff who cooperated and participated in the assessment process.

10. *Keep in mind that the primary emphasis in assessment is on the improvement of teaching, learning, and services to students.* If assessment results are not used to make program and curriculum improvements, the entire effort is wasted.

As noted above, this is a suggested list of implementation steps. Individual institutions may want to follow this sequence, add steps, change the order in which the steps take place, or delete a step that is not applicable.

Individual institutions may want to follow this sequence, add steps, change the order in which the steps take place, or delete a step that is not applicable.

Finally, I want to emphasize that there is no one perfect model or sequence of actions that fits all two-year colleges. Each institution should design a model and course of action to address its specific needs and circumstances.

Assessment Update
March-April 1995
Volume 5, Number 2

JEFFREY A. SEYBERT

Assessment of the Transfer Function

In an earlier column I outlined four possible conceptual models on which an institution might base an effort to assess institutional effectiveness. I would like to illustrate how one of those models (namely that proposed by the League for Innovation in the Community College) can be applied to assess a primary community college mission: the transfer function.

The transfer function or mission is historically the most traditional role of community colleges and, at the same time, one that has generated considerable recent attention and controversy. The transfer mission was the major reason for the establishment of the first junior colleges, and, while modern comprehensive community colleges have taken on numerous additional roles, transfer remains as a primary function. In fact, the last few years have seen calls from several prominent critics of the breadth of community college missions for a return to the pre-eminence of the transfer function, even to the exclusion of other mission components (for example, career/occupational education, remedial/developmental education, economic development, noncredit/community education, and the like). In any case, given the importance of the transfer function in modern community colleges, it seems appropriate to outline a framework to assess its effectiveness.

The transfer function or mission is historically the most traditional role of community colleges.

The League for Innovation assessment approach consists of three steps: identification of the major components of a college's mission, articulation of key or seminal questions (outcomes) for each component, and identification of appropriate measures to answer each question or assess each outcome. Since we have already identified the transfer function as the mission component we wish to assess, we can move to articulation of seminal questions regarding transfer. I will suggest and comment on four such questions.

1. *Do transfer students accomplish their community college education objective?* Since I have previously discussed the importance of satisfaction of students' educational objectives as a primary index of effectiveness for community colleges (see *Community College Strategies* in *Assessment Update*, 1992, 4(6), 15), I will not restate those arguments here. It is

Assessment Update
July-August 1993
Volume 5, Number 4

important to know, however, whether students achieved their objectives in several areas, including general education, their area of specialization or major, and a number of noncognitive areas.

2. *How do transfer students evaluate community college experiences and services?* Feedback from transfer students regarding the quality of instruction they received, their evaluations of student and support services, and their experiences with facilities and equipment can provide important information regarding transfer effectiveness.

3. *Do transfer students actually transfer?* If so, where and at what rates?

4. *Do transfer students succeed at senior institutions?* Success can be measured in terms of academic progress (for example, credit hours earned and rate of progress toward the baccalaureate degree), academic performance (for example, GPA overall and in the major), and the percent who actually receive a bachelor's degree.

These questions can be answered by using the specific assessment procedures that follow.

Feedback from transfer students can provide important information regarding transfer effectiveness.

Follow-Up Surveys of Transfer Students (Questions 1 and 2)

Follow-up surveys of transfer students conducted shortly after transfer (six months to one year) and after longer intervals (three to four years) provide valuable data regarding former students' perceptions of the degree to which they met their educational objectives; their growth in a variety of noncognitive, affective areas; and their evaluations of their two- and four-year college experiences and environments. Many community colleges have implemented transfer follow-up survey procedures or similar methodologies designed to gather opinion data from alumni or graduates.

Transcripts and Reports from Receiving Institutions (Questions 3 and 4)

Systematic exchange of student data between senior institutions and community colleges provides important information regarding overall transfer rates and former students' academic progress and performance. In a few cases (such as Illinois), such exchanges have been established on a statewide basis. There are also several instances in which groups of two- and four-year institutions have formally agreed to exchange such data on a regular basis, usually annually (for example, in Kansas, Texas, and California).

In Kansas, two community colleges (Johnson County and Kansas City, Kansas), nearby four-year colleges, two of the state's public research universities (the University of Kansas and Kansas State University), and several of the state's regional public universities have agreed on a set of 13 transfer student data elements that the universities and four-year colleges will provide

Assessment Update
July-August 1993
Volume 5, Number 4

to the community colleges annually. These data are combined with results of transfer follow-up surveys and data from the students' community college academic records to produce comprehensive transfer databases maintained by the community colleges. From these databases, the two-year colleges produce annual transfer follow-up reports for each participating senior institution detailing information regarding that institution's transfer student data, as well as an overall report covering all students who transferred from the community college in a given year. An effort is currently underway to expand this arrangement to include all 19 community colleges and six public universities in the state.

Assessments of Cognitive Outcomes (Question 1)

Assessment of transfer student learning outcomes is particularly difficult, since many students transfer before they achieve any formal academic milestone at the community college. In many states, receipt of an associate's degree carries no functional meaning at the senior institution. What transfers are credit hours; the degree confers no formal academic status (such as automatic junior class standing). Thus many (and, in some states, most) transfer students fail to apply for community college graduation and, in fact, transfer prior to earning enough credits to graduate. The transfer database at Johnson County Community College reveals that students transfer at all points in their career and that, on average, students complete approximately 28 to 30 hours before transfer. Thus, questions of when and how to "capture" these students for assessment are major unresolved issues for many community colleges. In addition, while assessment of general education might be theoretically feasible for transfer students, assessment of the major is even more problematic, since most students take only one or two courses in their major prior to transfer. For these reasons, those community colleges that attempt to assess transfer students' learning outcomes may decide to limit those efforts to the few who graduate with associate's degrees, at least until procedures can be developed to systemically reach a wider range of those students (for example, entry-level cognitive assessment at the four-year institution).

Questions of when and how to "capture" transfer students for assessment are major unresolved issues for many community colleges.

Thus, assessment of the transfer function, in the form of answers to the four seminal questions regarding transfer posed earlier, can be accomplished with the use of three or four data-gathering procedures. And, while I do not want to minimize the effort required to plan and implement these procedures, they are feasible for institutions that wish (or need) to commit the resources necessary to do so.

Assessment of Career and Occupational Education

Career students' achievement of their educational objectives is a primary index of a college's effectiveness.

In the two previous *Community College Strategies* columns, I outlined four possible conceptual models from which an institution might choose a basis for its assessment plan. Then I applied one of those models (the League for Innovation model) to assessment of the transfer function. I extend that discussion here to include assessment of the career and occupational education component of the two-year college mission. The career component plays a prominent role for many two-year institutions, encompassing career training, occupational retraining, continuing education courses, and degree programs. Thus, it is very important that these colleges include assessment of the career component of their mission in any overall effort to measure student outcomes and institutional effectiveness.

The league for innovation model involves three steps: identification of the component or components of the college mission to be assessed; articulation of key questions (outcomes) for that component; and identification of appropriate measures to answer the identified questions (that is, assess each outcome). Five such questions may be posed for career and occupational education:

1. *Do career students accomplish their community college educational objectives?* As is the case for transfer students, career students' achievement of their educational objectives is a primary index of a college's effectiveness. This is true for general education, the student's major, and in the affective domain as well.

2. *How do career students evaluate community college experiences and services?* Feedback regarding quality of instruction, perceptions of student and support services, and experiences with facilities and equipment provide useful input to inform an overall effectiveness assessment effort.

3. *Do career students obtain appropriate employment?* Since the *raison d'etre* for career education is training for high quality jobs with career advancement potential, we need to know whether students obtained employment in the

Assessment Update
November-December 1993
Volume 5, Number 6

BEST COPY AVAILABLE

types of positions for which they were trained, their degree of satisfaction with those positions, and, at some point, what their salary and promotion potential may be.

4. *How do employers evaluate career students' training?* Employers are in a unique position to evaluate career students' preparation in terms of work-related technical skills, communication skills, problem-solving skills, and interpersonal and teamwork skills.

5. *Is the college meeting local or regional labor market and economic development needs?* A major purpose of career education is to assist in meeting local or regional labor market demands for a highly trained, skilled technical workforce. In addition, many colleges see economic development in the form of contract training for business and industry and other initiatives as related to their career education mission.

These questions can be answered with four specific assessment methodologies:

Career students are a more stable and less transient student population than are transfer students.

Follow-Up Surveys of Former Career Students (Questions 1, 2, and 3)

Career student follow-up surveys conducted shortly after students complete their programs (for example, six months to one year) and after longer intervals (for example, three to four years) provide important information regarding the degree to which they met their educational objectives, their growth in a variety of cognitive and noncognitive areas, their evaluations of their college experiences and environments, and the appropriateness of and satisfaction with various aspects of their jobs. An important methodological issue is the availability of completion status and current demographic information (names, addresses, telephone numbers, and so on) of former career students. In some states (for example, Oregon), these data are maintained on a statewide basis by the state department of labor or employment, thus permitting the systematic exchange of information. In most cases, however, careful records must be established and routinely updated to allow for systematic surveying of former career students.

Assessments of Cognitive Outcomes (Question 1)

As I noted in the discussion of transfer, assessment of student learning outcomes for two-year colleges is a very difficult proposition. It is, however, somewhat less difficult for career students since they are a more stable and less transient student population than are transfer students. A greater percentage enroll in and complete specific programs and curricula and thus represent a viable population on which assessments may be conducted.

A variety of assessment methodologies is available to evaluate student learning, both discipline-specific and in general education. These include nationally standardized tests, locally produced instruments, portfolios of student work, capstone courses and experiences, internships in business and industry, final major projects, and others. Assessments of student learning have been in place in community colleges for many years in certain areas such as the allied health disciplines (for example, standardized licensing exams in nursing, dental hygiene, respiratory therapy, and others) and commercial art (portfolio evaluation).

Employer Surveys (Question 4)

Feedback from employers regarding career students' training provides a crucial source of information about program and instructional effectiveness. Any of a number of methodologies may be employed to gather this information, including focus groups of local employers and advisory committee members and surveys of employers of former students using names and addresses supplied by the students themselves or from statewide databases.

Assessment of career education can be accomplished using a relatively small number of straightforward methodological approaches.

Constituent Surveys and Economic Impact Studies (Question 5)

Effectiveness in meeting local labor market demands and economic development needs may be assessed in two major ways. First, surveys of constituents can determine whether employment needs are being met and the level of local residents' and employers' satisfaction with the college's efforts in those areas. Second, economic impact studies can demonstrate, at least in part, the college's tangible economic contribution to the local community and region.

As is the case for the transfer function, assessment of career education can be accomplished using a relatively small number of straightforward methodological approaches. And while the complexities of these procedures—in particular, assessment of student learning outcomes—must not be minimized, they are feasible for institutions willing to commit the resources to carry them out.

Assessment Update
November-December 1993
Volume 5, Number 6

JEFFREY A. SEYBERT

Assessment of General Education

Clearly there are two important facets of assessment of student learning outcomes: assessment in the major or discipline and assessment of general education. In this column I would like to make some comments about the latter. For community colleges, particularly in some transfer areas, assessment of student learning outcomes may, in fact, be limited to general education since in many "transfer majors" students take only one or two courses at the lower division (that is, two-year college) level. Thus there may be very little specific content acquisition to assess in any overall sense. In fact, many two-year colleges do not offer transfer majors per se; rather they award a generic associate of arts degree for students who wish to transfer. In this sense, then, assessment of general education takes on added importance for community colleges.

A first principle for general education is that it should be "general." That is, unlike discipline-specific curricula, neither its teaching nor its assessment should be the responsibility of individual departments or faculties. Presumably, such general types of knowledge as numeracy, communication, ethics, problem solving, cultural awareness, and others that often appear in statements of general education should be taught "across the curriculum" and thus are the responsibility of the faculty as a whole. Therefore, determination of appropriate outcomes and assessment methodologies in these and similar areas should be undertaken by multidisciplinary faculty teams rather than by faculty in individual departments.

For example, if communication skills are identified as an area of emphasis for general education, then, presumably, such skills are emphasized across the curriculum in a wide variety of courses. It follows that faculty from a variety of disciplines should be involved in determining appropriate outcomes and assessment methodologies for those skills. It should not fall to the English and speech faculties to make these decisions alone. Similarly, the results of general education assessments must be shared by the faculty as a whole. If those results are positive, everyone gets to take some of the credit. Likewise, if the results point out areas of weakness,

Results of general education assessments must be shared by the faculty as a whole. If those results are positive, everyone gets to take some of the credit.

Assessment Update
March-April 1994
Volume 6, Number 2

everyone shares the responsibility and should be involved in curriculum redesign and modification to produce improvement in those areas.

I would now like to suggest a series of steps that an institution might take in an effort to assess general education.

1. *Select a general education assessment task force or committee.* As I have emphasized previously in regard to assessment in general, this task force should be made up primarily of faculty, although a few carefully selected administrative/support staff (from the institutional research, testing and assessment, or instructional dean's offices, for example) may be of valuable assistance to the faculty committee. Assessment of general education, like all assessment of student learning outcomes, must, in my opinion, be faculty driven if it is going to be successful. Although administrative support and assistance are usually required, faculty should play a major role in any assessment effort, particularly to increase the likelihood that assessment results will be used to improve teaching and learning—the principal reason to do assessment in the first place.

Assessment of general education must be faculty driven if it is going to be successful.

2. *Identify the relevant general education requirements.* This step will not be as easy as it sounds. In many institutions, two or more statements of general education exist and may be in conflict. At my institution, for example, the outcomes assessment committee found a relatively short, somewhat vague statement of general education in the college catalog and a much longer, more detailed statement of the "aims of general education" in the college curriculum handbook. Thus an important initial task for the committee was to decide which of these statements would form the basis for the college's effort to assess the outcomes of general education.

3. *Divide the committee into multidisciplinary teams.* General education statements or requirements will almost always lend themselves to grouping in a relatively small number of categories (such as communication, math, or problem solving), and in fact many are already written this way. An effective way to begin the assessment process is to divide the committee into small multidisciplinary teams that correspond to these natural groupings and whose responsibility it will be to identify the appropriate outcomes and concomitant assessment methodologies.

4. *Identify appropriate outcomes.* In this stage of the process the multidisciplinary teams must translate often vague general education statements or sets of courses into specific outcomes. It is important to concentrate on identification, definition, and articulation of the outcomes and not to be concerned about assessment methodologies. In addition, the committee must identify the group of students who should have mastered these outcomes and for whom some assessment of outcomes is appropriate. This step is not as easy as it might sound.

5. *Identify or create appropriate assessment methodologies.* At this point in the process the committee, often with the assistance of institutional research and/or testing and assessment personnel, must find, select, or

Assessment Update
March-April 1994
Volume 6, Number 2

create measurement methodologies to assess the outcomes specified in step four. It may also be necessary to modify some of the outcomes during this step, particularly if they are found to be "unassessable." Adjustments and modifications are frequently necessary during the implementation of assessment. Assessment will always be a dynamic, evolutionary process because of changes in external mandates and other external forces, changes in internal conditions such as personnel and resources, and advancements in measurement methodologies and techniques. Change is a fact of life, particularly in a field as new as assessment, and the assessment process must be flexible enough to accommodate that change.

6. *Implement the assessment process.* Implementation of assessment will require commitment of both fiscal and human resources; it is not free! Administrators may have to find the resources necessary to implement assessment. Most important, assessment requires faculty involvement, particularly where resources are limited and must be reallocated to accomplish the goals of the overall assessment effort.

Finally, I have a few thoughts about students and assessment. Assessment should be as unintrusive as possible for students, particularly in community colleges. Since assessment occurs relatively late in students' careers at the institution, they may realize little if any direct benefit from the results of those assessments. It is not as realistic to make assessment a graduation requirement in a two-year college as it is in four-year colleges and universities. Given that an associate's degree has little importance for many transfer students, since it confers no official standing at transfer institutions in some states, making assessment a graduation requirement could drive down the already abysmally low graduation rates typically seen in community colleges. Similarly, appeals to participate in long, involved assessment processes to benefit "future generations" of students are not likely to generate student enthusiasm. In any case, any selection methodology that relies essentially on volunteers will be highly suspect in terms of the validity and representativeness of its results. Thus "alternative" or "authentic" assessment techniques, such as capstone courses and experiences, evaluation of portfolios compiled throughout the students' careers, and imbedding of assessment items in appropriate course final exams, may be more effective and feasible than mass standardized testing.

Assessment should be as unintrusive as possible for students, particularly in community colleges.

I believe that assessment of student learning outcomes should be independent of evaluations of student achievement for purposes of grading and satisfaction of graduation requirements. If students successfully complete course and graduation requirements, then they have earned the grades and degree specified by those requirements regardless of their performance on any assessment procedures. I consider the ultimate purpose of assessment to be improvement of curriculum, not evaluation of student performance. Thus, although students may be evaluated for a grade on "authentic" aspects of their performance, through portfolios and capstone experiences or courses, use of these experiences for assessment should be conducted separately, perhaps even after the students have left the institution.

Assessment Update
March-April 1994
Volume 6, Number 2

Assessing Student Learning

> ***My approach to assessment of student learning outcomes is not connected to evaluation of individual student performance.***

Several of my recent columns have been inspired by questions that are frequently asked of me as I travel around the country to talk with community college faculty and administrators about assessment. One such question has to do with ways to assess student learning. Given the growing agreement among faculty that course grades are not, by themselves, sufficient evidence of student learning, many of our colleagues interested in, or charged with, assessment responsibilities are at a loss as to how to proceed with assessment. This column is meant to be a brief, though incomplete, summary of ways to assess student learning outcomes. It does not include, for example, such legitimate methodologies as follow-up surveys, which are useful for determining the success of important programmatic and curriculum outcomes such as program completion and appropriate transfer or job placement rates.

My approach to assessment of student learning outcomes is not connected to evaluation of individual student performance for purposes of grading and describing completion and graduation status. The purpose of outcomes assessment, as I conceive it, is to evaluate the curriculum and the ways in which a program is delivered. To do this, we must somehow obtain evidence of performance from students. For purposes of outcomes assessment, I believe that this evidence must be compiled and analyzed in group form and not connected to any individual student. In other words, students should be evaluated individually based on the degree to which they complete the course and program requirements we have established, not on how well the curriculum is or is not doing what we say it is supposed to do, which is what assessment is all about.

Standardized tests are devised to assess either achievement in general education overall or some aspect or component of it, or achievement in a specific major or discipline. These instruments have presumably been normed on national or regional samples and their statistical reliability and validity established and published. In reality, the extent to which these instruments have been subjected to standardizing and norming procedures

Assessment Update
July-August 1994
Volume 6, Number 4

varies widely. Thus, any instrument under consideration should be carefully examined to make sure that it, in fact, meets any claims made for it by authors, publishers, and marketers. Standardized tests may meet considerable faculty resistance, particularly if the instrument under consideration does not closely reflect the curriculum in the area to be assessed.

State, regional, or national licensure examinations are examples of standardized tests and can be ready-made assessments for those disciplines for which they are available (for example, nursing, dental hygiene, and accounting). With the exception of applications in the few fields for which licensure examinations are required, standardized testing may be one of the most difficult assessment methodologies for use in community colleges due to serious problems in student identification, "capture," and motivation.

Locally produced tests are also knowledge-based but are developed by faculty on a given campus, often without psychometric standardization and norming. These assessments generally are not beset by many of the faculty objections raised against standardized tests, but they may be seriously flawed due to their lack of psychometric validation. Gary Pike has addressed these issues in several *Assessment Update* columns. Locally produced assessments may be "stand-alone" instruments or less formal sets of items embedded in some other in-place assessment and evaluation technique such as a course final. This type of unobtrusive assessment may be one of the most effective approaches for community colleges.

Standardized testing may be one of the most difficult assessment methodologies for use in community colleges.

Portfolios are collections of students' work, typically selected from throughout their academic careers to show growth or progression. Portfolios should be evaluated holistically by knowledgeable individuals, usually faculty members in the appropriate department. This evaluation often occurs after students have left or graduated and is not connected in any way to their grades or eligibility to graduate. Inherent to the use of portfolios are issues of which work samples to select, who selects them, and who maintains the portfolios. These considerations, as well as those surrounding portfolio evaluation, make this technique time-consuming, labor-intensive, and, therefore, expensive. It is, however, widely believed to also be one of the more effective and useful assessment methodologies.

Final projects are usually assigned or expected at or near the end of a student's academic career and are designed to reflect overall mastery of a discipline, set of disciplines, general education competencies (for example, problem solving, integration, analysis, synthesis, and communication skills), or some combination thereof. They may be graduation requirements and, as such, are graded and used as indices of students' eligibility for graduation. But their use as assessment tools should take place after and independently of their use as summative evaluations of individual student performance.

Capstone experiences and courses are similar to final projects but are more structured and formalized; they may be an entire course, portion of a course, or related experience such as an internship or work placement. Once again, such an experience may be part of a formal course, program, or graduation

Assessment Update
July-August 1994
Volume 6, Number 4

requirement and, as such, will be graded. However, its use as an assessment method must be separate from its use to evaluate student performance.

Capstone experiences and courses require students to demonstrate all or a portion of the skills and abilities they are supposed to have acquired as a part of their matriculation in and through a given program or curriculum. These can include both discipline-specific and more general competencies (for example, communication skills, ability to work and get along with a variety of individuals, and independent problem solving). As with portfolios and final projects, capstone experiences and their products, if any, should be evaluated by an expert or group of experts, preferably after their use as evaluations of individual student performance. The experts may be practitioners rather than, or in addition to, faculty; such assessment is a good way to use advisory committee members if they are already associated with a program.

This list of methods is not sufficiently detailed to be useful in designing a specific assessment process. I have simply provided a brief overview of some of the more important methodologies currently in use to assess student learning outcomes.

Assessment Update
July-August 1994
Volume 6, Number 4

JEFFREY A. SEYBERT

Assessment in Washington State Community and Technical Colleges

Guest Columnist: William S. Moore, assessment coordinator for the Washington State Board for Community and Technical Colleges.

In October 1989, the Washington Higher Education Coordinating Board endorsed a Student Outcomes Plan developed by the State Board for Community College Education. The goals of the plan were to assess the results of students' participation in a community college education, to identify strengths and weaknesses, and to improve the quality of the educational experience for students attending community colleges. Initially, the legislature provided funding only for system-level assessment activities. Later, the 1990 legislature provided supplemental funding for each community college to undertake local institutional assessment activities. Provision of process and data support to 27 community colleges as they developed and implemented institutional assessment plans has been a significant addition in scope and depth of the system effort. In spring 1994, funding was approved for incorporating the five technical colleges (merged into the system in 1992-1993) into the overall assessment initiative, beginning with 1994-1995.

The goals of the plan were to assess the results of students' participation in a community college education, to identify strengths and weaknesses, and to improve the quality of the educational experience for students attending community colleges.

Description of Current Activities

The assessment effort for the Washington Community College System has shifted in some significant ways since the beginning of the formal statewide initiative in 1989. The state board assessment activities began with system-wide research studies and data collection, addressing significant questions for each of the major mission areas of the community college system: technical education, transfer education, and basic skills education. The assumption was that system-level assessment projects would each lead to findings and results that would, in turn, move the system toward policy recommendations and improvements. We have learned that while such system-level research is valuable in addressing external accountability concerns, the relationship between assessment information and educational improvements is not that straightforward, particularly for the system as a

Assessment Update
November-December 1994
Volume 6, Number 6

BEST COPY AVAILABLE

whole. High quality system research does not lead quickly, or even necessarily, to significant policy implications or improvements for the system. We have discovered that we need to consider the array of assessment information we compile as a whole rather than as isolated studies, and we need to spend quite a bit more time than we had originally anticipated developing ways to engage college faculty and administrators in making judgments about assessment findings. For the system as a whole, we can incorporate what we learn into our overall budget process and explore the policies we control directly for their impact on our self-identified strengths and weaknesses.

Research Studies. The system's student outcomes research effort has provided valuable information about the strengths and weaknesses of the system's educational achievements. Although much has been learned, further work is in progress that will help to fill out the picture of the impact of community and technical college education. The following is a description of research work completed or in process during 1993-1994: we have studied students' status after leaving college, using the data match with the Washington State unemployment insurance file to provide employment data on former students for the past four years. The process is fully automated and an ongoing part of the system's operations. Colleges can use this outcomes information locally for program improvement and budgeting.

We need to consider the array of assessment information we compile as a whole rather than as isolated studies, and we need to spend more time developing ways to engage college faculty and administrators in making judgments about assessment findings.

Over the past several years, we have published several reports of study findings related to student and employer satisfaction. The Work Force Training and Education Coordinating Board has been given a mandate to coordinate a common system of student and employer satisfaction studies for secondary and postsecondary work force training providers, and we are working with that board to design a common approach. A survey of former academic transfer students that focused on student satisfaction with community college instruction and services and on experiences with the transfer process was conducted in 1992. Survey responses have been analyzed and incorporated into a report reviewing the range of transfer-related findings to date. Two reports have been produced summarizing student progress results from several data sources, including a pilot study conducted in 1992 at twelve colleges that followed the progress of former developmental education students in college-level programs. Computer programming required for automating this process is scheduled to be completed during fall 1994 so that colleges can follow the progress of their own developmental education students.

Systemwide Projects. The Student Voices Project is a collaborative pilot research project involving four community colleges and two four-year institutions-Seattle Central, Edmonds, South Puget Sound, Bellevue, Evergreen State College, and the University of Washington. The goal of the study is to gain a deeper understanding of how students evaluate their learning experiences, and to provide colleges with a model for further qualitative studies on other student populations.

Assessment Update
November-December 1994
Volume 6, Number 6

The Writing/Thinking Project has collected materials (assignments and student writing) and conducted retreats around assignment and assessment issues for four disciplinary areas in 1993-1994: nursing and allied health, fine and performing arts, natural sciences, and sociology and anthropology. Current plans include developing focused follow-up workshops on the evaluation of writing in specific content areas and compiling a resource handbook on effective assignments and assessments in the disciplines for which retreats have been held.

The Math/Quantitative Skills Project, a statewide effort to identify and promote innovative improvements in the teaching and assessing of quantitative reasoning skills, conducted a second assessment task retreat during 1993-1994 and has been active in maintaining connections both to the ongoing public education reform effort (Commission on Student Learning) and to assessment issues related to the development of the new technical degree being pilot-tested.

The Data and Technical Analysis Group has been organized to provide more extensive technical support and training for colleges interested in accessing and using existing data sources regarding student and institutional assessment. The State Board for Community and Technical Colleges has sponsored working retreats for this network of college technical resource people and produced reports on the use of the data base information for assessment and planning purposes and for college accreditation self-studies.

We have concluded that more attention needs to be focused on underachieving groups of students and on student persistence and success.

System-Level Uses of Assessment Information

Disseminating research results and provoking campus-level questions. Research findings are reviewed by a system group convened to advise each study. Research results have also been presented to state board members and system groups of college presidents, trustees, administrators, and faculty for review and discussion. These discussions have been the principal means for raising awareness about the system's strengths and areas in need of improvement.

Identifying research and policy implications. The various reports and findings have been reviewed and discussed by both SBCTC and college staff to determine what issues need further study. We have concluded, among other things, that more attention needs to be focused on underachieving groups of students and on student persistence and success.

Developing budget priorities. System-level outcomes assessment results have been used to establish improvement goals for the 1993-1995 operating budget request. The draft goals were reviewed by a wide range of system groups and were endorsed by the presidents and adopted by the state board. Assessment results are again being used for 1995-1997 biennium budget planning.

Assessment Update
November-December 1994
Volume 6, Number 6

Connections Between Assessment and Campus Reform and Improvement Efforts

Building campus-level assessment resources and expertise. While the system research has provided a broad view of student outcomes for community colleges as a group, the findings do not necessarily apply to each college. Once the research methodology was established as valuable, computer programming was implemented for the common student information system so that colleges can obtain the same information for their own students. We are also expanding our effort to facilitate activities designed to train college technical staff on ways to access and analyze these data to provide decision-making support for campus planning.

Intercampus information sharing about assessment projects and encouragement of quality assessment work. By producing the quarterly *Washington Assessment Group Newsletter* (currently mailed to about 1,800 people), by organizing the annual statewide assessment conference, and by compiling and distributing a computerized data base library on all major outcomes assessment-funded activities at the colleges, we are encouraging intercampus contacts about activities and potential improvement approaches in areas of common concern.

> *We are also expanding our effort to train college technical staff on ways to access and analyze these data to provide decision-making support for campus planning.*

Promoting faculty-driven program- and course-level improvements in key areas. In specific high-priority areas such as writing and mathematics and quantitative skills, we have taken an active leadership role in holding meetings for interested faculty representatives from the community colleges and four-year institutions. Our goals have been (1) to have the participants learn from one another particular strategies for teaching and assessing these areas and take these ideas back to their campus colleagues and (2) to compile and create resource documents—such as the *Quantitative Skills Assessment Task Report*—for wide distribution among the campuses. We plan to extend this approach in the coming biennium to other key areas such as competence-based technical education, critical thinking, and cultural diversity.

For more information about resources or reports cited in this article, contact William S. Moore, State Board for Community and Technical Colleges, 319 7th Ave., Olympia, WA 98504-2495. Tel.: (206) 586-8296. Fax: (206) 586-6440. E-mail: <wmoore@carson.u.washington.edu>

Assessment Update
November-December 1994
Volume 6, Number 6

JEFFREY A. SEYBERT

Assessment of Noncredit Continuing Education and Community Service Programs and Courses

This column is, in part, based on the research of John Cosgrove, St. Louis Community College, and Craig Clagett, Prince Georges Community College.

Continuing education and community service courses and programs offered on a noncredit basis constitute an important component of the overall mission of most two-year colleges. Such offerings include courses and programs for general interest and personal growth and development that range from martial arts to bass fishing, from microcomputer use to quilting, and from retirement planning to Chinese cooking. Many community colleges also provide noncredit training and retraining, sometimes on a contract basis to local business and industry and sometimes as a service to individuals in various professions to enable them to retain or upgrade their skills and expertise, often for the purpose of maintaining their professional licensure or certification status. It is also not uncommon for community colleges to schedule lectures, cultural arts performances, and other special events that are open to members of the community as a public service. Finally, two-year colleges often provide a wide array of community services to their constituents on a noncredit basis, including adult literacy and basic education classes, personal counseling, career planning and placement services, and information and referral services.

> ***It is somewhat surprising that so little emphasis has been placed on assessing the effectiveness of noncredit programs and courses.***

Given the pervasiveness of these noncredit programs and courses, it is somewhat surprising that so little emphasis has been placed on assessing their effectiveness. There are probably several reasons for this lack of attention to assessment of continuing education offerings. In many institutions, the continuing education and community service operation is required to be financially self-sufficient. Therefore, programs that either do not generate sufficient revenue to offset expenses or are sparsely attended are often not offered again. Thus, a type of natural selection operates in which repeat offerings become, in a sense, an index of a program's or course's effectiveness. A second possible reason has to do with the perception of noncredit continuing education programs in the overall postsecondary education setting. These programs are widely seen as being on the periphery of the

Assessment Update
March-April 1995
Volume 7, Number 2

education enterprise, as somehow second class, and thus not worthy of the attention paid to more traditional credit programs. In addition, many noncredit courses are offered at off-campus sites, and students or participants may be in contact with the program or course on a short-term basis, perhaps for only a single session of an hour or two. Thus, there is very little opportunity to apply the traditional data collection methodologies on which so much of the assessment of credit programs is based.

Community colleges involved in noncredit programming can undertake a series of steps to initiate an assessment effort. The first is the all important identification of the intended outcomes of noncredit programs and courses. This may involve the creation or revision of department or unit mission statements, careful consideration of course and program objectives, and articulation of specific outcomes that lend themselves to some type of measurement. Course and program offerings need to be examined closely to determine whether and the degree to which their curricula and objectives fit with the identified outcomes. And, finally, assessment instruments and methodologies and indices of effectiveness need to be identified to gauge the success of the programming in producing these outcomes. Such methodologies should take into account and be sensitive to the transitory nature of most student encounters with noncredit and community service programming. Examples of such indices and methodologies are outlined below.

Course and program offerings need to be examined closely to determine whether and the degree to which their curricula and objectives fit with the identified outcomes.

Some of the traditional indirect indices of noncredit continuing education effectiveness, which have been used informally for many years, are still appropriate. These include the "survival of the fittest" criterion for repeat programming mentioned above. If an institution offers programs and courses that employers and students persist in taking, then those programs and courses must be providing needed information in some effective way. Repeat business is a reliable sign of effectiveness. In addition, another traditional, desired outcome of noncredit programming, often unspoken, is the recruitment of noncredit students into credit courses and programs. Noncredit-to-credit migration is relatively easy to assess if reasonable data are collected and maintained regarding a few demographic characteristics of enrolled credit and noncredit students.

One very valuable and relatively straightforward assessment technique involves the use of noncredit student course evaluation forms that delineate and specifically evaluate the degree to which intended course and program outcomes are accomplished. Such forms must necessarily be customized for each course. However, experience at Johnson County Community College indicates that this approach can be pursued with preprinted scannable forms designed to provide space for later addition of specific course objectives and instructors' names.

Assessment Update
March-April 1995
Volume 7, Number 2

Another valuable technique is the use of student, participant, and employer follow-up evaluation surveys. Similar to those used to gather information from credit students and their employers, these instruments can provide important insights into the efficacy of noncredit programming. Follow-up surveys may be particularly useful for one-time events such as lectures and cultural arts performances. These surveys require some method of obtaining attendees' names, addresses, and, possibly, telephone numbers; however, if tickets for the event are required, this information may be available as a result of the ticketing process.

A similar technique involves the use of focus groups to gain evaluation information regarding noncredit programming. Focus group procedures may also be helpful in assessing the effectiveness of one-time educational and cultural events, as well as more traditional continuing education courses. In addition, information from focus groups can also be valuable in informing future programming. In other words, these methodologies can serve as needs assessments to guide future program planning and decision making at the same time they are being used to assess the effectiveness of past programs.

In all of these cases (student course evaluations, follow-up surveys, and focus groups), it is important to collect information regarding achievement of student or participant personal goals for the noncredit course, program, or event under study. As is the case for credit courses and programs, satisfaction of students' educational objectives is a primary index of community college effectiveness. It is preferable to elicit students' personal goals prior to their participation, and then assess their perceptions of the degree to which their goals were accomplished. Realistically, however, noncredit students or participants are often in contact with the institution for such a short time that asking them to report this information after the fact may be the only practical way to collect data. On the other hand, a one-minute paper may be a very useful way for noncredit instructors to acquire data about their students' goals for a course at its outset. These goals can then be compared, either in the aggregate or by individual student, with similar data gathered at the end of the course.

As is the case for credit courses and programs, satisfaction of students' educational objectives is a primary index of community college effectiveness.

Licensure renewal rates for individuals involved in mandated continuing professional education (nurses, dental hygienists, other allied health professionals, accountants, real estate professionals, attorneys, insurance agents, and others) may be another important measure of the success of noncredit programs. These data may be readily available in some cases but not in others. Those involved in assessment may ask appropriate agencies about the availability of this information.

Finally, growing numbers of noncredit continuing education programs have the goal of producing increases in economic development. Here, assessment of success in promoting economic growth and benefit to the service area or region is essential. Economic impact studies, preferably of the noncredit operation alone, but at least of the institution as a whole, can help make the case for success in relation to this goal.

Assessment of noncredit continuing education and community service programs has been a neglected component of most community colleges' efforts to evaluate and demonstrate their effectiveness. However, given the increasingly important role these operations have in overall college missions, it is important that we begin to assess the effectiveness of these operations.

Assessment Update
March-April 1995
Volume 7, Number 2

Community College System-Level Approaches to Core Indicators of Effectiveness

Measuring institutional effectiveness can be regarded as determining whether or not systems are doing what they say they are doing.

Guest Columnist: Loretta Seppanen, manager for research and analysis, Washington State Board for Community and Technical Colleges, Olympia.

Over the past several years there has been increasing interest in what are variously referred to as performance indicators, indicators of effectiveness, and core indicators. An interesting question arises, then, as to the extent to which state offices for community and technical colleges are involved in analyzing and publishing effectiveness indicators for the public two-year sector. That was the question addressed in a winter 1995 survey of state community college offices conducted by the Washington State Board for Community and Technical Colleges. Responses were received from 26 states. In light of the responses, it is fair to say that reporting of indicators at the system level is commanding considerable attention. Among the 26 respondents, 12 had published documents or reports describing the system outcomes on at least several core indicators, 8 were currently developing a report, and 6 had no plans for such a document.

What do I mean by reporting indicators of institutional effectiveness? Ewell and Jones (1994, p. 6) defined indicators as "policy relevant statistics produced regularly to support overall policy planning and monitoring at the national, state or system level." Measuring institutional effectiveness can be regarded as determining whether or not systems are doing what they say they are doing. All of these reports provide indicators of institutional effectiveness based on this definition, but each achieves that goal in its own way.

Differences among states in the indicators chosen stem in part from different purposes. Five different purposes emerged from a review of the narratives in the reports:

Complying. Several states are under mandates to report indicators. In some cases, these mandates were imposed and indicators specified in legislation. In other cases, states are anticipating such mandates and have created their reporting approaches as a way of directing the nature of the mandate or precluding mandates altogether.

Assessment Update
November-December 1995
Volume 7, Number 6

Persuading. Many states have a goal of shaping public or stakeholder opinion of the strengths of the two-year system. Expressions such as "meets the increasing demand from the public that higher education report on its accomplishments" illustrate this goal.

Monitoring. In some cases, states have defined standards for performance or goals for improvement. The reports serve to monitor progress on these standards or goals.

Improving. Most states claim that their reports will lead to institutional improvements. An illustrative comment is that the report will "help colleges focus on matters that are most indicative of student success." Essentially each report is designed to serve a formative evaluation function.

Judging. Some states see the reports as a way to make judgments about the quality of one institution versus another.

Responses from the following states revealed that they are not working on indicator documents: Idaho, Kansas, Michigan (which requires extensive Carl Perkins Standards and Measures reporting), Mississippi, New Jersey, Virginia, and Wyoming. One explanation given for not planning a state-level document was that reporting on core indicators is the responsibility of each college, not the state system. Some states have so few community colleges that a state report seems redundant to individual college reports. The response from Pennsylvania revealed that this state is just beginning work on a data base for indicator reporting.

It is not surprising that the types of indicators used vary significantly by state.

Among the states with published reports, most provide data for the system only. There were, however, exceptions to that pattern, as shown in Table 1. Of those in the planning stage, most states plan to publish data by individual college. Most respondents reported that they are disseminating or will report trend data. Some produced single-point-in-time data for the initial reports and will add trend data in the future.

Given the differences in the purposes of reports and differences in the state data systems, it is not surprising that the types of indicators used vary significantly by state as well. Six types of indicators are evident (see Table 2).

Inputs. Generally, states reported on inputs to address the issue of equal access. Some states also reported participation rates and other input indicators.

Process. Indicators include the number participating in exemplary programs or specialized accreditation. Some states reported faculty opinions about the institutions or studies of time-on-task.

Outcomes. Job placement, transfer data, progress from developmental to college-level coursework, and graduation data are the most common outcomes measures.

Contextual. Ewell and Jones (1994) called these resource efficiencies and effectiveness indicators. They include funding per full-time equivalent and availability of facilities. Other frameworks would include "inputs" as a contextual variable.

Table 1. Planned or Current Indicator Reports by Purpose and State

State	*Complying*	*Persuading*	*Monitoring*	*Improving*	*Judging*	*By System or College?*
Alabama[a]				X		College
Arizona[a]		X		X		College
California	X	X	X			System
Colorado[a]		X		X		College
Florida	X		X	X		System
Illinois[b]		X	X	X	X	College
Kentucky[c]	X	X				System
Minnesota		X		X		College
Missouri				X		College
New Hampshire				X		College
New Jersey[a]	X	X		X		System[d]
New Mexico[a]		X		X	X	College
New York		X				System
North Carolina	X			X		College
Oklahoma				X		College
Oregon		X				System
Texas[a]	X	X	X	X	X	System
Washington[a]		X				System
Wisconsin[a]				X	X	College

[a] State has a framework for a report, but no report has been issued.

[b] Series of reports rather than one document are planned as data are analyzed.

[c] Kentucky reported that Tennessee has a similar report though Tennessee did not respond to requests for a copy.

[d] State law requires that each college issue a report.

Table 2. Planned or Current Types of Indicators by State

State	*Inputs*	*Process*	*Outcomes*	*Contextual*	*Return on Investment*	*Customer Satisfaction*
Arizona	X		X		X	
California	X		X	X		
Colorado	X		X			X
Florida	X		X			
Illinois	X	X	X	X	X	X
Kentucky	X	X	X	X		X
Minnesota	X	X	X	X		X
New Jersey	X	X	X	X		X
New Mexico			X			
New York	X	X	X	X		
North Carolina	X	X	X			
Oklahoma	X		X			
Oregon	X		X			
Texas	X		X	X		X
Washington	X		X	X		

Note: No information was available for planned reports in Alabama, Missouri, New Hampshire, or Wisconsin.

Assessment Update
November-December 1995
Volume 7, Number 6

Return on investment. This requires an analysis of the net benefit of participating in two-year colleges as compared to the cost. Only two states proposed a return-on-investment indicator.

Customer satisfaction. States are using surveys of students and employers for this indicator.

Most states used the American Association of Community Colleges (1994) report *Community Colleges: Core Indicators of Effectiveness* as a guide for at least some of their indicators. Colorado will be developing indicators directly from that report. Several states found the guide not applicable to the system-level analysis, and several were not acquainted with the report. Most states have designed their indicator report based on data that are currently available. Some, however, listed desired indicators for which there are no current data.

Two examples illustrate the extent of diversity among the states. Oregon publishes a simple 2-page report featuring seven outcomes indicators related to the major mission areas. The report provides data on five of the indicators and a promise of data on the other two in the future. By contrast, the 72-page New York document lists hundreds of indicators, about half of which are contextual in nature: median family income, state appropriations, and so on. The report also includes reader-friendly trend-line graphs.

Considerable staff energy is being devoted to developing or reporting system-level effectiveness indicators in more than half of the states.

In summary, considerable staff energy is being devoted to developing or reporting system-level effectiveness indicators in more than half of the states. Some 75% of the responding states are investing time and expertise in the task of reporting on system effectiveness indicators. These efforts serve diverse ends in the various states. The most common purposes are, however, persuading stakeholders of the quality of the two-year sector and improving the system. There is, as well, considerable diversity regarding indicators, though outcomes indicators are common to every report and input indicators are nearly so.

The author invites readers to share additional information: Loretta Seppanen, Washington State Board for Community and Technical Colleges, P.O. Box 42495, Olympia, WA 98504-2495. Tel.: (360) 753-3685. E-mail: <Seppanen_Loretta/SBCTC@ctc.edu>

References

American Association of Community Colleges. Community Colleges: *Core Indicators of Effectiveness*. Washington, D. C.: American Association of Community Colleges, 1994.

Ewell, P. T., and Jones, D. P. "Pointing the Way: Indicators as Policy Tools in Higher Education." In S. S. Ruppert (ed.), *Charting Higher Education: A Source Book on State-Level Performance Indicators*. Denver: Education Commission of the States, 1994.

Assessment Update
November-December 1995
Volume 7, Number 6

Virginia Community College System

Guest Columnist: Debra Banks, coordinating producer of assessment teleconferences and owner of D. L. Banks Associates, a consulting firm specializing in planning, research, and assessment for both for-profit and nonprofit organizations.

Assessment may require knowledge and technical skills not usually possessed by faculty and administrators charged with its design and implementation.

Assessment may require knowledge and technical skills not usually possessed by faculty and administrators charged with its design and implementation. Thus, considerable professional development may be required to help those individuals attain the requisite skills. These development activities usually include attendance at conferences, review of relevant literature, and, possibly, use of consultants. Virginia Community College System (VCCS), however, has recently begun to use distance learning technologies to provide information about assessment to faculty and staff throughout the state. Specifically, in fall 1995, the VCCS Assessment Group presented workshops on assessment design and techniques via interactive video teleconferencing.

In 1994, VCCS initiated teleconferencing as a way to provide professional development opportunities efficiently and effectively for larger audiences than are usually reached by more traditional methods. The purpose of the teleconferences is to promote professional development through the sharing of innovative ideas and the enhancement of knowledge and skills.

The focus of the assessment teleconferences emerged from a survey conducted by the VCCS Assessment Group. Responses indicated that less than half of the group's members had formal training in research design or assessment and evaluation techniques. In addition, the members expressed a need to learn more about valid and reliable assessment design, data analysis, and reporting.

Further surveying helped establish the outline of topics presented in the conferences. This outline was transformed into two interactive discussion and question-and-answer sessions facilitated by John Muffo, director of academic assessment at Virginia Polytechnic Institute and State University,

Assessment Update
March-April 1996
Volume 8, Number 2

BEST COPY AVAILABLE

and Karen Gentemann, director of institutional assessment at George Mason University. The sessions were broadcast live to the colleges and universities in Virginia, and viewers were asked to call in their questions.

The teleconferences focused on several critical areas of the assessment process, including assessment design, identification of the specific assessment target or problem and its parameters or boundary conditions, understanding of who requested the assessment information and data and how this information would be used, and selection of appropriate assessment methodologies. Also discussed were the validity and reliability of methodologies and instruments, details of various assessment methodologies such as surveys and standardized tests, sampling procedures, and effective reporting and report formats.

These assessment teleconferences offered several advantages in providing professional development across a wide range of institutions. First, production costs of the conferences were less than $4,000. Expenses for individuals attending a conventional conference in person would have far exceeded this figure. Second, broadcasting to individuals' work sites created the potential for much broader participation and relieved them of the task of dealing with the details and problems associated with transportation to and housing at a conference site. Third, the receiving sites videotaped the conferences for the convenience of individuals who were unable to view the live broadcasts. Thus, teleconferences are not time and place bound, as conventional conferences are. Fourth, the presenters became more widely recognized as expert resources who could be called on later by faculty and administrators who might have additional questions. Finally, several viewers commented that they planned to use the videotapes as training tools for professional development in program and classroom assessment in their institutions.

Given the obvious advantages in using distance learning technologies, the VCCS assessment teleconference approach appears to be a viable methodology that can easily be used in other settings.

Given the obvious advantages in using distance learning technologies, continued budget constraints in higher education, and increased demand for professional development for assessment (and, in fact, a wide variety of other areas), the VCCS assessment teleconference approach appears to be a viable methodology that can easily be used in other settings. However, it is important to note that such an endeavor requires a skilled and experienced production staff; presenter-facilitators who are content experts, who have knowledge of audience needs and level of sophistication, and who are relatively comfortable in front of the camera; and a coordinator-producer to facilitate the production staging and to oversee public announcements of the sessions.

Copies of the VCCS assessment teleconference videotapes can be obtained at cost from the VCCS office. Requests should be directed to Linda Carr, Director of Educational Planning, VCCS, James Monroe Building, 101 N. 14th Street, Richmond, VA 23219. E-mail: <carrl@so.cc.va.us>.

Assessment at Two-Year Institutions

A biennial national survey on the state-of-the-art in assessment practices is conducted.

Assessment Update
July-August 1996
Volume 8, Number 4

Guest Columnists: Michael K. Smith, director of the Clearinghouse for Higher Education Assessment Instruments at the University of Tennessee, Knoxville; and Jama L. Bradley, coordinator for undergraduate services for the Department of Communication at Georgia State University, Atlanta.

The Clearinghouse for Higher Education Assessment Instruments, located at the University of Tennessee, Knoxville, conducts a biennial national survey on the state-of-the-art in assessment practices. The intent of this survey is to determine how widespread assessment practices are at colleges and universities throughout the United States. The survey also attempts to determine the types of outcomes that are being measured and the types of methods or techniques being utilized.

Survey results are compared and contrasted across various types of institutions: two-year, four-year, and comprehensive institutions, and universities. Here we focus on results related to two-year institutions, drawn from both the survey in 1994 and the previous survey conducted in 1992. Surveys were mailed to chief academic officers. In 1994, over 1,000 institutions responded to the survey, including approximately 400 two-year institutions.

In 1994, 42% of survey respondents reported that they had campuswide assessment programs; another 54% noted that they expected their assessment programs to become campuswide in the near future.

Twenty-one percent of the respondents noted that they had an assessment program more than three years old; this was an increase from 10% in the 1992 survey. Thirty-seven percent of institutions reported having programs in their early stages (one to three years old); 17% were beginning assessment activities; and 17% were just starting to think about assessment. It appears that assessment activities are becoming more widespread among two-year institutions, although many are still in the planning or initial stages of development.

BEST COPY AVAILABLE

The survey queried respondents on the measurement of seven outcome domains: basic skills, general education, major field, critical thinking, student satisfaction, college environment, and affective development. For each domain, respondents were asked which of the following measurement techniques they utilized: commercially developed instruments, faculty-developed instruments, computer adaptive test, opinion survey, projects (including papers), portfolios, and interview methods. Respondents could check more than one technique.

In 1994, 82% of two-year colleges reported using commercial instruments to measure basic skills, compared to 74% in 1992. In contrast, only 54% of four-year institutions, 58% of comprehensives, and 49% of universities assessed basic skills with commercial instruments. Faculty-developed instruments to measure basic skills were used by only 35% of two-year institutions, while nearly half of four-year institutions, comprehensives, and universities noted that they had developed their own instruments.

In 1994, 33% of two-year institutions used commercial instruments to assess general education (national average was 34%), and 20% used faculty-developed instruments (national average was 21%). These percentages are roughly equivalent to those in the 1992 survey, indicating that the assessment of general education is not prevalent and not increasing.

Campuswide assessment activities at two-year institutions are widespread and continue to increase.

The widest variety of techniques was used to assess knowledge in the major field or technical career. In 1994, commercial instruments were in use at 21% of two-year institutions, faculty-developed instruments at 26%, opinion surveys at 14%, projects at 15%, portfolios at 14%, and interviews at 9%. These percentages represent only slight increases over those reported in 1992. The same diversity of methods was found in other institutions, but with higher frequency. For instance, 31% of four-year institutions, 45% of comprehensives, and 37% of universities reported using portfolios in major fields in the 1994 survey.

Commercial instruments were the main vehicle used to measure critical thinking (17% of two-year institutions used such instruments). Student satisfaction was primarily, and frequently, measured with opinion surveys (63% of such usage in 1994); college environment was also primarily assessed via opinion surveys (31% in 1994); and affective development was measured with opinion surveys infrequently (15% in 1994). These results represent only slight increases over the 1992 findings. Two-year institutions were very similar to other institutions in assessment of these domains.

Campuswide assessment activities at two-year institutions are widespread and continue to increase. Data from these surveys indicate that the primary domains assessed are basic skills, major or technical field, and student satisfaction. The use of multiple measures seems to be widespread only in measuring a student's knowledge in the major or technical field. There is certainly room for growth in most institutions' assessment programs.

A complete copy of these survey results can be obtained for $15 from the Clearinghouse for Higher Education Assessment Instruments, 312 Claxton Education Building, University of Tennessee, Knoxville, TN 37996.

Assessment Update
July-August 1996
Volume 8, Number 4

Developing a Statewide Core Indicator Project for Community Colleges

CCCOES has directed most attention to career and technical education outcomes from the onset, especially measures of graduate placement.

Guest Columnist: Richard A. Voorhees, associate vice president for Educational Support Services at Colorado Community College and Occupational Education System.

Twelve state-supported community colleges form the postsecondary component of the Colorado Community College and Occupational Education System (CCCOES). Annually, these colleges serve over 90,000 students, marking CCCOES as one of the largest community college systems in the nation. The system also serves as the state oversight agency for career and technical education, including area vocational schools, local district colleges, and secondary vocational education. CCCOES has a long history of accountability initiatives, particularly in the career and technical arena. External developments in this current decade have accelerated CCCOES's interest in accountability, resulting in the development of statewide core indicators of effectiveness for its colleges. The context in which these indicators were developed and the experience gained in implementing this effort are the topics of this article.

CCCOES has directed most attention to career and technical education outcomes from the onset, especially measures of graduate placement. This economic connection should not be surprising, given mandates for hard data found in federal vocational legislation. In response, CCCOES developed a comprehensive unit record vocational data base in 1976, permitting extensive longitudinal tracking of career and technical graduates. Over 200,000 student records enter this system annually. Enhancements to this data base were necessitated by the Carl Perkins Applied Technology and Vocational Education Act of 1992, which requires states to document student learning outcomes in basic skills, general occupational skills, and specific occupational skills. Efforts to comply by vocational and instructional administrators and faculty on the community college, area vocational school, and secondary levels have increased awareness of accountability.

Other precursors to the CCCOES core indicators of effectiveness include

Assessment Update
March-April 1997
Volume 9, Number 2

BEST COPY AVAILABLE

Colorado statute HB 1187, enacted in 1987. This law requires each public postsecondary institution to produce annual accountability reports for the Colorado Commission on Higher Education (CCHE). Initially, this process produced the effect intended by the Colorado legislature, ensuring that higher education was approaching accountability seriously. This process gradually lost its luster in the eyes of legislative leadership as required institutional reports were judged to be useless for purposes of assessing statewide performance. Concomitantly, national agencies such as the Education Commission of the States and the American Association of Community Colleges (AACC) issued increasingly frequent calls to more clearly document institutional outcomes.

CCCOES's presidents, instructional and research personnel, and system staff began to discuss the pros and cons of developing a set of common performance measures late in 1994. Utilizing AACC's (1994) report on core indicators of effectiveness as a framework, CCCOES's staff began working with campuses and presidents to investigate the feasibility of implementing indicators across the system. Early discussions focused on balancing the need to be responsive to accountability driven by the community college mission with the need to educate the public about the context within which community colleges operate. Concerns expressed early in these discussions included the potential image problems engendered by publishing a student persistence rate that would be significantly lower than that of four-year institutions.

Student goal attainment became the most necessary ingredient in the CCCOES core indicator of effectiveness effort, but also the most difficult to apprehend in Colorado and elsewhere.

The AACC report suggested thirteen core indicators intended to be responsive to the community college mission: student goal attainment, persistence (fall to fall), degree completion rates, placement rate in the work force, employer assessment of students, number and rate of transfer students, performance after transfer, success in subsequent coursework for basic skills students, demonstration of critical literacy skills, demonstration of citizenship skills, client assessment of programs and services, responsiveness to community needs, and participation rate in service area. The strong recommendation that persistence rates not be published in isolation, but rather in combination with degree completion rates, placement rates in the work force, and transfer rates, allayed concerns that any single indicator, by itself, might assume preeminence. Consensus for proceeding with the project developed around the need to assess students' success on the basis of their goals. Given the multiplicity of student goals and their likelihood of changing with exposure to the institution, student goal attainment became the most necessary ingredient in the CCCOES core indicator of effectiveness effort, but also the most difficult to apprehend in Colorado and elsewhere. Changes to the student application form and a query routine built into the telephone registration system have improved the system's ability to capture student goals, and ultimately to compare their stated goals with their actual attainment. Such measurement is longitudinal in nature and at this writing awaits completion in the CCCOES core indicator framework.

Assessment Update
March-April 1997
Volume 9, Number 2

Most indicators could be captured by CCOES's student information systems, including the unit record vocational data base and a statewide unit record data system maintained by CCHE. The latter source was particularly helpful in tracking transfers from their native community colleges to other colleges and determining their subsequent academic performance at receiving institutions. Census data for each college's service area were tabulated and compared with enrollment to calculate participation rates. Colleges were asked to supplement these systemwide data-gathering activities with employer assessments of graduates, demonstration of critical literacy and citizenship skills, client satisfaction with programs and services, and responsiveness to community needs. Unlike the quantitative measures developed at the system level, colleges were encouraged to be creative in constructing multiple measures of each of these indicators. A first report on the CCCOES core indicators of effectiveness was published in December 1995 and subsequently delivered at the Association for Institutional Research National Forum in Albuquerque in May 1996 (Voorhees and Cheng, 1996).

Lawmakers' frustration with the status quo and wide interest in statewide performance measures resulted in HB 1187's repeal in the 1996 legislative session. In its place is new legislation requiring CCHE to develop a set of performance indicators for higher education. At this writing a statewide technical committee is culling a set of indicators for initial review from a wide set proposed by the National Center for Higher Education Management Systems. The current strategy is to select a set of seven or eight indicators for all public institutions to implement, leaving seven or eight for each institution, or governing board, to develop. This scheme promises to allow for institutional buy-in. Colleges can develop indicators to extend or supplement the CCHE-required indicators. The experience that CCCOES and its constituent colleges gained through developing Colorado's first statewide core indicator project promises to be invaluable as a new era of accountability unfolds.

The current strategy is to select a set of seven or eight indicators for all public institutions to implement, leaving seven or eight for each institution, or governing board, to develop.

References

American Association of Community Colleges. *Community Colleges: Core Indicators of Effectiveness.* Washington, D.C.: American Association of Community Colleges, 1994.

Voorhees, R. A., and Cheng, D. "Core Indicators of Effectiveness: Colorado Community College and Occupational Education System." Paper presented at the Association for Institutional Research National Forum, Albuquerque, New Mexico, May 1996.

Assessment Update
March-April 1997
Volume 9, Number 2

A New Paradigm for Evaluating Transfer Success

Guest Columnists: Michael Quanty and Richard Dixon of Thomas Nelson Community College; Dennis Ridley of Christopher Newport University.

The key to improving transfer is faculty involvement.

Thomas Nelson Community College (TNCC) and Christopher Newport University (CNU) have been trying to improve transfer success rates for years. In 1989, with support from the State Council for Higher Education in Virginia's Funds for Excellence program, we set out to establish a new standard in transfer research. Reasoning that the key to improving transfer is faculty involvement, we reassigned time for two faculty members from each college to review data, design follow-up studies, and then make recommendations to improve the process. For a full year the faculty and research staffs pored over a variety of analyses and a plethora of data. Using detailed records for more than 1,800 students who transferred, we created regression analyses predicting grade-point average at CNU; profiled successful and unsuccessful transfer students; and explored the effects of hours attempted and completed at TNCC, degree type, courses taken, and every conceivable demographic variable. We also employed a consultant to interview students regarding their experience at the colleges.

At the end of the year, we had uncovered a number of statistically significant findings. However, when it came to making recommendations we were at a loss. Even though we had examined students' success from every angle we could imagine and spent hours discussing the data, we had no idea what should be done. Knowing that women and older students experience less "transfer shock" provides little basis for action. Faculty cannot change students' demographic backgrounds. They also cannot make students complete at least thirty hours before transferring.

In the aftermath of this frustration, however, we experienced an epiphany; we made a paradigm shift. The result is the Course-Based Model of Transfer Success (CBMTS), a new model for evaluating transfer that we maintain can greatly enhance the ability of the nation's community colleges and four-year colleges and universities to improve the success of transfer

Assessment Update
March-April 1998
Volume 10, Number 2

BEST COPY AVAILABLE

students. It provides faculty with action-oriented results they can use to improve students' preparation. The basic difference between this new model and those commonly employed in transfer research is that it is course-based rather than student-based.

This shift is simple but crucial. Whereas traditional research tracks particular students from the community college to their transfer institution, the CBMTS yields information that shows how well students who complete course prerequisites at a community college perform in specific *courses* compared to *students* who complete the prerequisites at the receiving college. The emphasis is on how well courses prepare students. This pinpoints for faculty exactly where students experience difficulty. It also creates a sense of urgency. Faculty agree that students who pass their course are prepared for subsequent courses which require that course as a prerequisite.

CBMTS is simple and relatively easy to implement, and provides comprehensive data that are immediately and obviously relevant to faculty at both the sending and receiving college. In short, it provides useful information. We developed a tracking system that examines every course having a prerequisite that could be met at CNU, TNCC, or at another college. For each course so identified, the program provides a grade distribution for students, broken out by semester and by whether the prerequisite was taken at CNU, at TNCC, or at another institution. A summary for each course totals grades across all semesters, and a discipline summary totals grades of all courses in the discipline and then for all semesters included. Chi-square analysis determines whether any observed differences are statistically significant. We now have data from spring 1990 through spring 1997 for TNCC and CNU.

The real strength of our new model is that when a problem is identified it can be attributed to a specific course at the community college and at the receiving college.

The results we have obtained thus far are very encouraging for community colleges. Generally, we have found that students who complete course prerequisites at TNCC perform at a level at least equivalent to students who complete prerequisites at CNU. The real strength of our new model is that when a problem is identified it can be attributed to a specific course at the community college and at the receiving college. Faculty take ownership for students who have successfully completed their course(s). If those students are not prepared for subsequent coursework, faculty want to know why. Demographic considerations do not matter when a faculty member has certified that the student has mastered course requirements.

This paradigm also works well within a total quality or customer service orientation. The receiving institution and the transferring institution have a customer-supplier partnership with shared interest in students' success. Identifying course-specific deficiencies allows quick and continuous improvements to be implemented and tested.

In October 1996 the Fund for the Improvement of Postsecondary Education (FIPSE) awarded us a grant to develop a "generic" version of CBMTS that can be adapted to support other college partnerships. Currently we are working with most of the community colleges and four-year colleges and universities in Virginia to develop and test a program that will give

Assessment Update
March-April 1998
Volume 10, Number 2

course-based comparisons using data that are collected centrally by a state council (a number of other states have similar arrangements). The information required to run the model can be produced easily by most colleges even without centralized reporting. Partners need to be able to share three types of files: (1) course files that include a student identifier, course identifier, course grade, and term (one record per student per course); (2) a target-course file that lists all courses at the four-year college or university that have prerequisites; and (3) a course-equivalency file that identifies community college courses that transfer as prerequisites for particular target courses.

With these basic files and a commitment to improvement, colleges can begin examining transfer in an exciting new way. Our experience at both TNCC and CNU has shown that faculty respond positively to the information generated by the model. When a problem is found, faculty talk to one another and find a solution. We also have found the model useful in a number of other applications. CNU, for example, uses it as one means of assessing the effectiveness of courses delivered on-line. They determined that students who complete courses on-line perform comparably to traditional students in follow-on courses. In fact, once we saw the power of the model, we began to use it in many contexts. It works; it gives faculty data that are both useful and motivating.

When a problem is found, faculty talk to one another and find a solution.

Through FIPSE support, we will be helping other colleges adopt CBMTS and customize it to their specifications. If you would like to learn more about the model, we encourage you to contact one of us directly or to visit our Web site at www.cnu.edu/cbmts.

The Wisconsin Technical College System Institutional Effectiveness Model

Guest Columnists: Donald Bressler, vice president of occupational and academic services at Northeast Wisconsin Technical College; and Deborah Mahaffey, bureau director of student and support services at the Wisconsin Technical College System.

The challenge rested not just in identifying specific indicators but also in achieving consensus on a core of indicators that would guide systemwide and *institutional improvement efforts.*

Assessment Update
July-August 1998
Volume 10, Number 4

Context for the WTCS Model

Many colleges have developed a model for assessing institutional effectiveness (IE), but the Wisconsin Technical College System (WTCS) was among the first in the nation to develop a systemwide model. The challenge rested not just in identifying specific indicators but also in achieving consensus on a core of indicators that would guide systemwide *and* institutional improvement efforts. The WTCS presidents and the state governing agency agreed to create a model that concurrently provided a measure of effectiveness of the system as a whole while honoring the autonomous nature of each of the sixteen colleges in the statewide system. In order to accomplish the task of implementing an acceptable IE model with a meaningful outcome, two key elements were imperatives in the process: cross-functional collaboration and representative consensus.

The WTCS is governed by a board appointed by Wisconsin's governor. A director hired by the board heads a state agency that provides system leadership. Each technical college is directed by a president hired by a local board. The college presidents work together, along with the state agency, under a formal structure called the Administrators Association. The structure includes subcommittees for each of the administrative entities, including instructional services, student services, research and planning, and administrative/financial services. In addition, numerous systemwide meetings are conducted, manager-level committee meetings are held, and ad hoc groups are created to address timely issues, recommend uniform practices, and undertake collaborative efforts. Collaboration is a valued way of doing business.

BEST COPY AVAILABLE

The colleges have embraced a collaborative approach to program evaluation. The systemwide evaluation model set forth in the 1980s incorporated three phases: *accountability* through annual monitoring and screening data to indicate the general health of a program; *evaluation* through in-depth assessment of programs and services; and *impact appraisal* through assessment of "value-added" dimensions of the program or service. Further, the WTCS as a system has a long history of conducting employer satisfaction surveys and longitudinal studies of graduate employment status to assess program success.

Wisconsin Technical College System Institutional Effectiveness Model
Core Indicators

Student Achievement and Satisfaction	***Employer Satisfaction***
Identification of student needs, goals, and interests	Employer satisfaction with graduates' work skills and performance
Identification of student functional skills at entry	***Organizational Quality***
Course completion	Achievement of institutional goals and standards
Students' grades	
Students' satisfaction with courses, programs, and services	Organizational climate
Student retention and withdrawal rates	***Public Perception and Satisfaction***
Student completion and graduation rates	Articulation and linkages with external organizations
Student achievement of educational goal(s)	Identification of customer needs and expectations
Student knowledge and skills at exit	
Pass rates and scores of licensure exams	Public satisfaction
Placement rates and employment success	

Development of the Model

In 1992, the WTCS Administrators Association and the state agency agreed to contract cooperatively with outside consultants to help develop a tailor-made IE model. The project was undertaken as a systemwide effort to identify core indicators that could be used locally by colleges to respond to accountability initiatives, accreditation demands, federal reporting requirements, and quality management initiatives. The development process was intentionally designed to be inclusive and to achieve consensus across the technical college system.

> *The development process was intentionally designed to be inclusive and to achieve consensus across the technical college system.*

One of the first steps centered on creating a task force with representation from various subgroups of the Administrators Association, the state agency, and other key stakeholders, including faculty unions. It was the mission of the Institutional Effectiveness Task Force to provide guidance and direction for the development of the model. Inclusiveness was paramount to the process. The task force assembled more than twenty face-to-face focus groups to obtain input from primary stakeholders regarding (1) key institutional characteristics of performance outcomes that define effectiveness in the technical college system and (2) indicators and data-gathering methods recommended for measuring effectiveness.

Focus groups included college presidents; students; college board members; union groups; faculty and support staff; college vice presidents; administrators for instruction, student services, adult education, economic development, research and planning, administrative services, and marketing; and state agency staff. The focus group activities produced a list of more than one hundred potential effectiveness indicators. The next round of input, via written survey, sifted the indicators down to what were considered the primary IE indicators.

Ultimately, seventeen indicators focused on four categories of effectiveness: student achievement and satisfaction, employer satisfaction, organizational quality, and public perception and satisfaction. The exhibit below shows the specific effectiveness indicators associated with each category.

Working with the Model

A new cross-functional steering team was established to provide the systemwide leadership and coordination necessary for moving from developing the model to adopting and implementing it. Clearly, the first stage had been inclusive of all customers, both internal and external to the technical colleges; this stage was time-consuming and thorough, if not laborious. Yet no one questioned the need for building all the right connections. Buy-in and consensus building take time and were recognized to be essential for the model to be credible and viable.

Seventeen indicators focused on four categories of effectiveness: student achievement and satisfaction, employer satisfaction, organizational quality, and public perception and satisfaction.

Leadership from the state agency and three college administrative areas—instruction, student services, and research and planning—made up the Institutional Effectiveness Steering Team. The team placed emphasis on continuing the focus on the Institutional Effectiveness Model (IEM); establishing priorities based on identified college needs; ensuring integration of data-collection systems; facilitating collective improvements in accountability measures; and coordinating cooperative endeavors. The work began by focusing on student achievement and success—not an unreasonable choice given that each college faced an external accreditation mandate to develop a plan for assessing student achievement. It was easy to see the gains to be achieved through collaboration. The work was to be carried out through cross-functional work groups. Work group membership and responsibilities were structured to ensure success: (1) individual members interfaced with college planning processes to integrate the results of IE measures with institutional improvement efforts; (2) membership included key stakeholders—representatives from all colleges, various college functions, and the state agency; (3) there was a consistent process for presenting the charge and mission to each work group; (4) each work group used an effective process-improvement strategy; (5) there was interaction with other appropriate groups and initiatives with related functional responsibilities; and (6) each work group had a limited focus and scope of measurement.

Three work groups were created to focus on seven of the indicators relating to student assessment and achievement. Each of the three work groups (student

Assessment Update
July-August 1998
Volume 10, Number 4

goals at entry and exit; student functional skills at entry and exit; and course completion, retention, and graduation) were chaired by a steering team member and charged with the task of making recommendations for effective implementation of the indicators at the college level, the state level, or both.

During the next two years, more than fifty-five individuals from across the state worked collaboratively on cross-functional work groups. These groups refined definitions of effectiveness indicators, clarified parameters, and offered recommendations to operationalize indicators at the college and statewide levels. The outcomes have been impressive. For example, the cross-functional work group on student goals piloted an instrument and process for assessing student goals at entry and suggesting points of interaction based on student needs. The approach of the cross-functional work group dealing with assessment of students' functional skills at entrance and exit produced a philosophical grounding for assessment activities, a "menu of options" for assessment, and a database on current practices with the WTCS. The work group on student retention and withdrawal rates and course completion and graduation rates developed standard definitions for terms and core analysis methods for the purpose of monitoring and assessment. It is important to understand that before recommendations were advanced and considered as final work group products, individual members of each work group shared preliminary recommendations with their respective colleagues to gain consensus on direction and to generate buy-in for full adoption and use of the recommendations. Those actions were considered to be standard process steps and at the heart of the rollout process.

Before recommendations were advanced and considered as final work group products, individual members of each work group shared preliminary recommendations with their respective colleagues to gain consensus on direction.

Progress to Date

The accomplishments of the three initial work groups served as a good starting point and springboard for the IEM. In fall 1997, a WTCS assessment conference was held to serve as a celebration of accomplishments, a forum for sharing assessment practices among individual colleges, and an opportunity to report on the status of the development of the IEM. All sixteen colleges participated, and the conference evaluations called for "doing this again."

The Administrator's Association has called for development of a comprehensive plan for systemwide implementation of all IE indicators. The IE Steering Team has promulgated a final deployment plan endorsed by the Administrators Association that calls for alignment with current evaluation models and practices. Further, the plan continues to rely heavily on systemwide collaboration. Intact administrative committees and targeted cross-functional groups have been asked to review current practices and collaboratively recommend "best practice(s)" related to the other areas of the IEM: employer satisfaction, organizational quality, and public perception and satisfaction.

Clearly, the work is not done. But, most important, the work done to date has been based on collaboration and consensus. The commitment is there.

Assessment Update
July-August 1998
Volume 10, Number 4

The benefits have been significant and meaningful in terms of both product and process. The work will continue to be refined and adjusted to meet changing WTCS customer needs and expectations, and cross-functional collaboration and representative consensus will continue to be imperatives in the approach of the Wisconsin Technical College System.

The benefits have been significant and meaningful in terms of both product and process.

Assessment Update
July-August 1998
Volume 10, Number 4

Campus Profiles

Peter J. Gray

About the Author

Peter J. Gray received his Ph. D. in educational psychology from the University of Oregon and an MS degree in curriculum theory from Cornell University. Dr Gray's areas of expertise include research on higher education; higher education assessment; course, curriculum, and program evaluation; and institutional research. As Associate Director of the Center for Support of Teaching and Learning at Syracuse University his responsibilities include managing the design, implementation and reporting of evaluation, research, and instructional development projects. These projects are intended to improve academic and non-academic programs, to assess student learning, and to inform administrative policy setting and decision making. This entails working closely with faculty, staff, and administrators throughout the University. He also is responsible for the day-to-day administration of the Center and a staff of over 20 people. Dr. Gray holds adjunct associate professor appointments in the Syracuse University School of Education's Higher Education Department and the Instructional Design, Development, and Evaluation Department. In addition, he chairs the Measurement and Benchmarking Council of the Syracuse University Improving Quality (SUIQ) Steering Committee. With Trudy W. Banta, he co-edited *The Campus-Level Impact of Assessment: Progress, Problems, and Possibilities* (New Directions for Higher Education, no. 100, winter 1997). He was the editor of and a contributor to *Achieving Assessment Goals Using Evaluation Techniques* (New Directions for Higher Education, no. 67, fall 1989). He has been a contributing editor to *Assessment Update* since its inception. Dr. Gray served on the Board of Directors of the American Evaluation Association from 1987 to 1995 and he has consulted with faculty, educational administrators, and government officials in the United States, Indonesia, Russia, and Ukraine on topics related to the improvement of higher education. Dr. Gray has produced over 100 reports, professional presentations, and workshops.

Contents

Page

Campus Profiles 297

University of Illinois at Urbana-Champaign 303
King's College, Wilkes-Barre, Pennsylvania 306
Western Michigan University 309
Kean College of New Jersey 311
James Madison University 313
The Ohio State University's Center for Teaching Excellence 316
The Evergreen State College, Olympia, Washington 318
The Office of Educational Development at the University of California, Berkeley 320
Northeast Missouri State University 323
Miami-Dade Community College 326
Ohio University's Multidimensional Institutional Impact Assessment Program 330
Ball State University 334
State University of New York College at Fredonia 337
Kent State University Regional Campuses 341
Winthrop College, Rock Hill, South Carolina 344
Capital University 347
Midlands Technical College 350
Mt. Hood Community College 354
Union College 358
Eckerd College 361
Virginia Military Institute 364
University of Colorado at Boulder 368
Miami University, Oxford, Ohio 372
Chicago State University 376
Western Carolina University 380

Page

Sinclair Community College 384
St. Cloud State University 388
Indiana University Bloomington 393
Assessment of Student Learning and Development at Winthrop University: An Update 397
Truman State University 401
Broadening the Scope of Assessment at Ball State University 404
Student Outcomes Assessment at Ohio University 408
Institutional Effectiveness at Midlands Technical College 412

PETER J. GRAY

Campus Profiles

The *Campus Profiles* column has appeared in the *Assessment Update* publication since 1989. Over 30 institutions that have made serious attempts to design and implement assessment programs have been featured. These have included large and small schools, public and private institutions, community colleges, liberal arts colleges, and comprehensive and research universities.

As is clear from the profiles of institutions in this compendium, assessment is intended to fundamentally change campus cultures and the way teaching and learning occurs. The introduction of assessment often results in changes in curricula, courses, programs, and structures, with the ultimate purpose being the improvement of student learning and development. In the Ohio University profile (vol. 3, no. 5, 1991), President C. J. Ping is quoted as saying, "Assessment of quality must have consequences and must offer direction for decisionmaking and action, in order to be valuable to the university." In the Virginia Military Institute profile (vol. 6, no. 2, March-April 1994) it is noted that, in order for an assessment program to thrive and make a difference, assessment must be part of the planning-evaluation-resource allocation cycle. Statements such as these are examples of support by high level institutional leaders in the introduction of assessment on a campus. Such support also has come from state legislators, boards of trustees, presidents and provosts, and leaders of professional societies and accrediting bodies.

The case studies represented by the profiles in this compendium suggest that in order for assessment to be successfully institutionalized on a campus it must be made a priority over an extended period by central and departmental leaders (see Truman State profile, vol. 9, no. 5, September-October 1997). "Leaders ... are those persons or groups who can mobilize human, material, and symbolic resources toward specific ends. ... Mobilizing resources in any social system depends upon the ability of leaders to direct the behavior of others" (Rosen 1984 cited in Curry 1992). As Palmer (1993) suggests, "The most powerful kind of leadership is to offer people pathways and permissions to do things they want to do but feel unable to do for themselves. That sort

> ***"Assessment of quality must have consequences and must offer direction for decisionmaking and action, in order to be valuable to the university."***

BEST COPY AVAILABLE

of energy evokes energies within people that far exceed the powers of coercion" (p. 9). Such leadership taps into people's intrinsic motivation for competence, success, quality, and continuous improvement. The profiles in this compendium illustrate the many different ways that educational leaders have motivated and supported assessment over extended periods (Gray, 1997).

One way that campus leaders have provided long-term support for assessment has been through the endorsement of guidelines, principles, or goals that set the institutional context for the development and implementation of assessment by local faculty and unit administrators. Many of the campus profiles in this compendium include such guidelines, for example, as Kean College (vol. 1, no.4, 1989), Evergreen State College (vol. 2, no. 3, 1990), Midlands Technical College (vol. 5, no. 2, March-April 1993), Chicago State University (vol. 6, no. 6, November-December 1994), Western Carolina University (vol. 7, no.2, March-April 1995), and Ball State University (vol.10, no.1, January-February, 1998).

Because assessment is intended to stimulate changes at all levels and in all areas of a higher education institution, it is best viewed as an innovation. As defined by Rogers (1995, p. 11), "An innovation is an idea, practice, or object perceived as new by an individual." Many times the introduction of an innovation like assessment is seen as an institutional event and the assumption is made that adoption will be accomplished just because a decisionmaker has announced it. However, as Hall and others (1975, p.52) have pointed out, "innovation adoption is a process not a decision-point—a process that each innovation user experiences individually." Therefore, they define change as a process not an event, which is undergone by individuals first, and then by institutions. The change process is a highly personal experience that entails developmental growth in feelings, skills, and knowledge.

The profiles in this compendium illustrate the many different ways that educational leaders have motivated and supported assessment over extended periods.

Although it may not have been formally recognized as such, in reading the campus profiles in this compendium, one can see the attention that has been paid to the idea of assessment as an innovation that can change a campus' culture. For example, at Truman State University (formerly, Northeast Missouri State University), the following stages of assessment were implemented: readiness, implementation, acceptance, and commitment (vol. 3, no.2, 1990). There were two phases of Ohio University's experience with assessment. In the first phase, university-wide assessment information was provided to faculty and staff and, in the second phase, individual academic departments were supported in their need for assessment information (vol. 10, no. 2, March-April 1998). As noted in the second profile on Truman State, in looking back over the history of assessment, we can clearly see that assessment information by itself was not sufficient for transforming an institution (vol. 9, no. 5, September-October 1997). Magruder, McManis, Young (1997) indicate that it was necessary to create a "campus culture that not only embraces assessment conceptually but actively supports organizational self-evaluation and the desire for improvement" (p. 7).

The retrospective profiles on Truman State (vol. 9, no. 5, September-October 1997), Ball State University (vol. 10, no. 1, January-February 1998), Ohio University (vol. 10, no. 2, March-April 1998), and Midlands Technical College (vol. 10, no. 5, September-October 1998) show the evolution of assessment as it becomes part of and transforms a campus culture.

At many institutions, there were structures and activities established to facilitate the gradual and systematic introduction of assessment into the campus culture. For example, campuses established assessment task forces or committees, offices and/or coordinators, student and faculty advisory boards, and grants or other funding programs. In addition, they sponsored events such as workshops and seminars, provided consultation and other technical support, and created and disseminated reports at different levels, from individual to institutional. All of these activities were intended to introduce assessment to the faculty, staff, and students of an institution in a personally relevant way.

As assessment has evolved on the campuses in this compendium, it has been defined in many different ways. At Midlands Technical College, for example, assessment involves planning (documenting the intended purpose, direction, and expected outcomes of the college and providing foresight in the formulation of policies, programs, and services) and evaluation (a process of measuring the college against its stated purpose and indicators of effectiveness in terms of the outcomes accomplished) (see vol. 5, no. 2, March-April 1993 and vol. 10, no. 5, September-October 1998). Wolf, at IU Bloomington (vol. 7, no. 6, November-December 1995), defined assessment as research and inquiry into the improvement of teaching and learning. In some cases states have mandated the operational definition of assessment as in Ohio where assessment must include incoming student characteristics, the nature and impact of educational experiences inside and outside the classroom, graduating student characteristics, and alumni evaluation of their experience. At Ball State University, in the early 1990's the focus of assessment shifted away from a value-added emphasis to a more comprehensive approach to assessing the educational environment and its impact on students (vol. 10, no. 1, January-February 1998).

Given the variety of definitions of assessment described in the compendium profiles, it is not surprising that the purposes of assessment were quite varied.

Given the variety of definitions of assessment described in the compendium profiles, it is not surprising that the purposes of assessment were quite varied. The following four purposes suggest the range included in the profiles:

1. to identify the strengths and weaknesses of the institution (mission; programs in general and general education; faculty; facilities, e.g., classrooms, labs., studios, residence halls, student centers, and gyms and recreational space; admissions; auxiliary services; registrar and registration operations; and computing)
2. to identify the strengths and weaknesses of a curriculum or program (congruence of the goals and objectives of the general education or a discipline-specific curriculum or program with individual course content, the sequence of courses, general instructional strategies, and the

cumulative impact of the educational experiences of the curriculum or program on learning)

3. to identify the strengths and weaknesses of course-related teaching and instructional elements (congruence of the goals and objectives of a course with course content, instructor behavior and characteristics, instructional methods and materials, student activities, examinations and other methods used to assess and grade students, and the impact of the course on learning)
4. to identify students' learning styles as well as their academic, personal and social interests, strengths and weaknesses (in order to provide academic and personal advising regarding programs of study and course selection, to provide remedial instruction, and to provide personal growth and development opportunities).

Such a wide variety of assessment definitions and purposes necessitates an assortment of assessment methods.

Such a wide variety of assessment definitions and purposes necessitates an assortment of assessment methods. The profiles also provide a rich source of examples of assessment methods that include:

- SWAT (Students at Winthrop Assessment Teams, vol. 4, no. 4, 1992), which focuses on issues of particular interest to students and on assessment projects that might be more effectively conducted by students;
- tracking through logs, student effort and commitment in and out of class;
- longitudinal, developmental portfolios of academic, personal, and social growth and development, as well as free writing exercises and individual student growth plans;
- faculty-conducted interviews of students;
- course-embedded assessment including placement testing and pre- and post-assessments of liberal arts and core curriculum areas/courses, as well as course-specific exams, quizzes, assignments, projects, and other classroom assessment techniques;
- mid-career (sophomore/junior) diagnostic activities such as an exam or project and capstone (senior) experience such as an integrative course, seminar, and/or project;
- regularly collected data such as enrollment and retention statistics of courses and programs (departments, schools, colleges), graduation rates, placement test scores, and SAT/ACT scores;
- national standardized tests (ACT-COMP, ETS Student Profile, GRE, NTE, MFAT, UAP)
- facilities assessment (lecture halls, standard classrooms, high tech classrooms, seminar rooms, laboratories, and studios); and
- alumni and employer surveys (e.g., one year out and every five years thereafter) of satisfaction with general education, programs of study, continuing education, and employment.

The profiles in this compendium document the great creativity and dedication that has gone into the development and implementation of assessment on these campuses. In a small way, as a group, these profiles document the impact that the assessment movement has had on higher education in the United States over the last 10 to 20 years. Ultimately this is the goal of assessment, to improve teaching and learning within courses and programs on each campus so that institutions can achieve their unique missions for educating their students. These profiles chronicle the many meaningful changes that have occurred for the people at the colleges and universities represented.

The messages that come across loud and clear from the profiles in this compendium are the importance of flexibility in developing locally meaningful assessment programs, the importance of strong and consistent leadership, and the importance of following a process of planned change. Through the involvement of external consultants, internal experts, faculty opinion leaders, and students, campuses have crafted assessment programs that uniquely suit their cultures and the purpose(s) of assessment that they have embraced. The profiles in this compendium provide examples of these efforts and they provide recognition of the people who helped to guide them.

These profiles document the impact that the assessment movement has had on higher education in the United States over the last 10 to 20 years.

References

Curry, B. K. Instituting enduring innovations: Achieving continuity of change in higher education. *ASHE-ERIC Higher Education Report, No. 7*. Washington, DC: The George Washington University, School of Education and Human Development, 1992.

Gray, P. J. Viewing assessment as an innovation: Leadership and the change process. In P. J. Gray and T. W. Banta (eds), *The Campus-Level Impact of Assessment*, New Directions for Higher Education, no. 100. San Francisco: Jossey-Bass, 1997.

Hall, G. E., Loucks, S. F., Rutherford, W. L. and Newlove, B. W. Levels of use of the innovation: A framework for analyzing innovation adoption. *The Journal of Teacher Education*, spring 1975.

Magruder, J. M., McManis, A., Young, C. "The Right Idea at the Right Time: Development of a Transformations Assessment Culture." In P.J. Gray and T.W. Banta (eds.), *The Campus-Level Impact of Assessment*, New Directions for Higher Education, no. 100. San Francisco: Jossey-Bass, 1997.

Palmer, P. J. Good talk about good teaching: Improving teaching through conversation and community. *Change*, November/December 1993.

Rogers, E. M. Diffusion of innovations. New York: The Free Press, 1995.

PETER J. GRAY

University of Illinois at Urbana-Champaign

Campus Profiles is intended to provide examples of the broad spectrum of assessment activities designed to improve institutions that are taking place in higher education. A major criterion for developing a profile is an institution's track record in following through on assessment findings.

Campus Profiles ***is intended to provide examples of the broad spectrum of assessment activities designed to improve institutions that are taking place in higher education.***

The institution profiled in this inaugural column is the University of Illinois at Urbana-Champaign. Much of the assessment and follow-up work at the University is the responsibility of the Office of Instructional and Management Services (IMS). IMS has for years been responsible for the development and administration of evaluation and assessment programs that systematically collect data to assist campus decisionmaking. The source of information for this column is a paper presented at the annual meeting of the American Educational Research Association (John Ory, "Evaluation and Assessment of Higher Education Decision-Making," April 1988) and recent conversations with John Ory, director of the office.

Structure. The Office of Instructional and Management Services (IMS) reports directly to the vice-chancellor for Academic Affairs. It has two divisions: instructional development and measurement, and evaluation.

Purpose. The range of assessment program activities includes decision research, policy analysis, and improvements in courses, curricula, and programs. Assessment for course improvement is intended to provide information to instructors for their use in making changes in individual courses. Assessment related to curriculum and program improvement is intended to provide information on the instructional and noninstructional aspects of academic, administrative, and other university units. Policy analysis is the systematic collection of data to inform institutional priority- and policy-setting. Decision research is assessment intended to provide information that is relevant to making specific decisions, such as whether a faculty member should be promoted or granted tenure, whether a student should be placed in one level of mathematics or another, or whether a new facility or program should be created. In some cases, decision research may

Assessment Update
Spring 1989
Volume 1, Number 1

BEST COPY AVAILABLE

be called "needs assessment" or "market research." Because such purposes as those just described are broad, they typically are related to sharply focused assessment programs that provide information relevant to various audiences.

Methodology. The Instructor and Course Evaluation System (ICES) is a computer-based system for obtaining student ratings. Faculty can select from a catalog of over 600 items classified by content area, such as instructional environment, instructor characteristics, and specificity. Certain items are always included: demographic items, global items regarding teaching effectiveness and course quality, and open-ended items; other items can be included at the discretion of the department and instructor. In addition, faculty can elect to include a set of items that is used throughout the university, with results reported in a student publication. The ICES is administered during the last two weeks of the semester. Results are presented in a computer-generated report that goes directly to the instructor. It includes descriptive statistics and a comparison of the global item ratings given to the instructor with the normative ratings of the instructor's department and of the university databases. Faculty have the option of sharing the results with their department administrators for the purposes of promotion and tenure.

The impact of the ICES has been to provide systematically collected information, which faculty feel ownership for, that is used for course improvement and for administrative decisions when so desired.

The Council on Program Evaluation (COPE) is a multi-faceted data collection program. A key element is the Program Evaluation Survey (PES), a paper-and-pencil instrument designed to assess students' perceptions of and satisfaction with the instructional, curricular, advising, and operational phases of academic departments. The survey includes biographical items to help identify various student populations and items to indicate, among other things, students' satisfaction with the various programmatic aspects of their departments. COPE also includes (1) faculty evaluation of the unit's administration, morale, operational procedures, and organizational context; (2) data on the national reputation of the unit; and (3) data concerning the unit's track record regarding tenure, promotion, courses taught, and budgets. The vice-chancellor for academic affairs chooses units for review on a cyclical basis. The Advanced Placement and Proficiency system (APP) is used to determine where new freshmen should be placed and whether proficiency credit and advanced placement should be offered. Using College Board Advanced Placement Program scores, departments establish policies for awarding proficiency credits and advanced placement for each score on the five-point APP scale.

The Transition Program Assessment uses four national measures, two locally developed instruments (Rhetoric Essay Exam and Mathematics), and group interviews to assess students in a high school transition summer program. These measures are administered prior to the first day of summer classes and at the end of each academic year to gather information about students' academic and personal growth, attitudes toward school, and assessment of the special program.

Assessment Update
Spring 1989
Volume 1, Number 1

Impact. The impact of the ICES has been to provide systematically collected information, which faculty feel ownership for, that is used for course improvement and for administrative decisions when so desired. The PES is used both as a part of a cyclical COPE assessment and as an ongoing measure of student perceptions and satisfaction. As a result, specific actions such as student meetings to improve communication, revisions in curricula, and improvements in academic advising and other support programs have been implemented. The impact of the APP has been very direct for the students involved. Nationally, the third highest number of Advanced Placement scores (4754) was submitted to the University from May 1987 candidates. The Transition Program Assessment impact has been to attract attention to the program's strengths and the need for improvement and to make appropriate changes in the students' orientation to the program.

Resources. It is not enough to allocate time, energy, and expertise to performing assessment; there also must be a commitment to carrying out a program of improvement once the implications of an assessment process are identified. In faculty-based systems like COPE, the talents of other university personnel are tapped for the purpose of conducting assessment and utilizing the results. In addition, units like audiovisual services and administrative computing provide essential support.

The Future. It is clear that a move is taking place from quantitative assessment to instructional development and faculty support. Other institutions should be aware that once a reliable, valid, and valued quantitative system is established, there will be considerable demand for assistance in implementing the results of assessment. The sooner such mechanisms can be instituted the better.

Contacts. For more information about assessment and its utilization at the University of Illinois at Urbana-Champaign contact: John C. Ory, Office of Instructional and Management Services, 307 Engineering Hall, 1308 West Green Street, Urbana, IL 61801, or Larry A. Braskamp, Office of the Dean, College of Applied Life Studies, 1206 South Fourth Street, Urbana, IL 61820.

It is not enough to allocate time, energy, and expertise to performing assessment; there also must be a commitment to carrying out a program of improvement once the implications of an assessment process are identified.

King's College, Wilkes-Barre, Pennsylvania

Instead of being viewed as merely a means of accountability, assessment is considered a way to improve both learning and teaching.

King's College is a Catholic liberal arts college with approximately 1,750 full-time students. Its course-embedded assessment program was started with external grant money that provided for course reductions during the academic year and summer stipends, in order to free faculty to engage in curriculum development. The first round of development served as a model for subsequent efforts, which were funded internally.

Background

In the last ten years, a series of planned changes have taken place at King's College. Before possible changes in the curriculum were even discussed, five years were spent on faculty development. Only after an outcomes-oriented curriculum had been agreed upon was the concept of assessment introduced. Slowing the pace gave faculty time to understand the changes needed for the college to attain its stated goal of excellence.

The curriculum development process resulted in a new core curriculum, consisting of three liberal learning areas: (1) transferable skills of liberal learning; (2) knowledge, traditional disciplines, and interdisciplinary perspective; and (3) responsible believing and acting. These areas define learning that is clearly more difficult to assess than learning in traditional areas, but faculty at King's College believe that new curriculum represents the real measure of quality in higher education.

The assessment component of the course-embedded model is defined as learning, rather than as measuring. Instead of being viewed as merely a means of accountability, assessment is considered a way to improve both learning and teaching. For example, assessment is used diagnostically, to identify students' strengths and weaknesses, and then prescriptively, to direct students to appropriate instructional experiences. Faculty also use assessment information to monitor the success of instruction and suggest changes in instructional content, methods, and materials. Assessment

Assessment Update
Summer 1989
Volume 1, Number 2

results are also used to address curricular issues such as sequencing; congruence among course objectives, teaching strategies, and examinations; and the cumulative impact of the educational experience on student learning.

Methodology

Students are assessed as part of the natural teaching/learning process in the classroom, while documentation of cumulative learning is provided. Assessment strategies are implemented in all core curriculum and major program courses. Students take assessments seriously because they count as part of the course grade, even though faculty also use assessments to provide information on the general accomplishment of learning objectives and to guide curricular improvement.

The course-embedded assessment model has six major components. The first is the use of pre- and post-assessments in all core curriculum courses and sections. Comprehensive growth plans that take place in the courses in a student's major program make up the second component. The third is the sophomore-junior diagnostic project that is begun in relation to a second-semester sophomore course in the student's major, continued over the summer and completed in a first-semester junior course, where a grade is assigned and credit awarded. The fourth component is senior-level integrated assessment that links learning in the core curriculum with learning in the student's major field, as part of a senior seminar. The fifth is the use of the ACT-COMP exam as a pre-assessment of freshmen and a subsequent, required post-assessment during the senior year. Alumni surveys, which occur one year after each student has graduated and every five years for all students, are the sixth component. Results are used to document outcomes over a more extended period.

Students take assessments seriously because they count as part of the course grade.

Impact

During the four years that the outcomes-oriented curriculum and course-embedded assessment program has been in operation, four dramatic changes have resulted: (1) the graduation rate has risen from 58% to 72%, (2) new records for the number of freshman applicants have been set during the last three years, (3) faculty indicate that students perform at higher levels than before, and (4) faculty are revitalized, actively seeking ways to improve their courses, curricula, and assessment.

Not all of the program elements were totally effective. Donald Farmer, academic vice-president and change agent for the program, states, "Not all experiments are equally successful; not all criteria for assessing student learning are equally specific; not all competence-growth plans are equally effective; not all senior-level integrated assessments are equally challenging. But all these are moving toward the successful implementation of an outcomes-oriented curriculum that uses course-embedded assessment strategies to help students realize their potential as learners." The proof of the program's success will be in the outcomes attained by the 1989 and future graduates, who will have spent all four years of their undergraduate education in the outcomes-oriented curriculum.

The Future

The faculty project teams that developed the curricula and the assessment strategies still meet regularly to monitor and refine them. In addition, a partnership is being formed between academic affairs and student affairs to develop a unified set of outcomes representing both areas. In this regard, a faculty member will spend half of his or her time facilitating further improvements.

The model is fully described in *Enhancing Student Learning: Emphasizing Essential Competencies in Academic Programs*, published by King's College (1988). For more information, contact Donald W. Farmer, vice-president and dean of academic affairs, or Edmund Napieralski, coordinator for core curriculum and assessment, King's College, Wilkes-Barre, PA 18711.

The faculty project teams that developed the curricula and the assessment strategies still meet regularly to monitor and refine them.

Assessment Update
Summer 1989
Volume 1, Number 2

PETER J. GRAY

Western Michigan University

Western Michigan University (WMU) is a regional institution with state and national responsibilities in instruction and research. WMU serves over 25,000 students through eight undergraduate colleges, a graduate college at one campus, and five regional centers around the state. The Office of University Assessment was established in 1987 within the Office of Academic Affairs, in cooperation with the Office of Institutional Research. The assessment effort depends on existing committees and councils for its design, its implementation, and responses to its results. This arrangement accords with the belief that assessment information will prove most useful to the extent that it fits with and meets the needs of the existing institutional structures.

George M. Dennison, provost and vice-president for academic affairs, and Mary Anne Bunda, director of the Office of University Assessment, note that the primary function of assessment at WMU is to facilitate improvement of academic programs and of the total educational environment. Recognition is provided to administrators who use assessment information to motivate faculty. Additional resources have been allocated to administrators who use assessment information as the basis for change.

In the study of general education, a standardized test was used to provide an external benchmark. The objective form of the ACT-COMP was chosen because it fits well with the university's definition of general education. In addition, skill subscore reports can be produced for use by faculty, as can individual reports that provide students with information concerning their growth.

When the ACT-COMP was first chosen, faculty members of appropriate curriculum committees were invited to help administer and sit for the test, so that they could become familiar with it and accept it more readily. Major curriculum policymakers were informed of the test's correspondence with WMU's general education program. The ACT-COMP also became a graduation requirement, since few seniors had volunteered to take the test. By providing students with developmental information, faculty with

The primary function of assessment at WMU is to facilitate improvement of academic programs and of the total educational environment.

Assessment Update
Fall 1989
Volume 1, Number 3

BEST COPY AVAILABLE

curriculum information, and policymakers with an external benchmark, WMU achieved acceptance of and involvement in its assessment program. WMU also has initiated dialogue with high schools about basic skills development and has designed new data collection systems appropriate for students' basic skills levels, especially where the effective use of verbal symbols is concerned.

Another component of the WMU program is an alumni survey, which the Office of University Assessment developed with the assistance of survey methodology experts. The classes of 1981 and 1986 were surveyed for their satisfaction with the institutional environment and academic programs. Departmental chairs, deans, and student services directors also provided information on alumni satisfaction. College curriculum committees have used the results to guide changes in instructional programs, and policy changes have been made regarding better coordination between academic programs and such support programs as academic advising.

Alumni were surveyed for their satisfaction with the institutional environment and academic programs.

Dennison and Bunda have identified three premises that guide their assessment efforts (in P. J. Gray, ed., *Achieving Assessment Goals Using Evaluation Techniques*, New Directions for Higher Education, no. 67, San Francisco: Jossey-Bass, 1989). First, it is essential to make the objectives of the assessment program clear. The institution must set student development as its top priority. Second, assessment must involve the faculty and other key decisionmakers from the outset, thereby securing their commitment to using the results. Third, assessment must be differentiated from related evaluative efforts, such as program review.

Assessment Update
Fall 1989
Volume 1, Number 3

PETER J. GRAY

Kean College of New Jersey

In 1986, Kean College of New Jersey implemented its Excellence and Equity Project, a plan for institutional change. Existing organizational elements were examined, and new elements were created and institutionalized, often replacing old ones. Kean's assessment program was developed to support the major objectives of this project.

Kean College has an enrollment of about 12,500 students. In the freshman class that entered in fall 1988, 61% were white, 18% black, 15% Hispanic, and 6% other minorities and foreign students. Most Kean students are from surrounding urban areas. Many are educationally and economically disadvantaged, the first members of their families to attend college. In addition, nearly all students commute to the campus. Kean has 40 undergraduate degree programs, ranging from arts and sciences to professional disciplines, with approximately 330 resident faculty.

Five assessment guidelines were developed by the president's task force and approved by the faculty senate, the board of trustees, and the administration.

Five assessment guidelines were developed by the president's task force and approved by the faculty senate, the board of trustees, and the administration. First, assessment is a means of exploring the curriculum and the learning associated with it, rather than an end in itself. Second, an assessment program should be consistent with the mission of a college, its programs, and its students' characteristics. Third, criterion-referenced instruments are preferred over norm-referenced ones, particularly in assessment of individual programs, and faculty have primary responsibility for determining the appropriateness of each. Fourth, assessment strategies adopted by faculty should be carefully selected, valid, reliable, and directly tied to the learning objectives of each course. Fifth, different assessment methods (licensing or other professional exams, faculty-developed measures, senior seminars, course-cluster exams) are appropriate for different programs.

Three Excellence and Equity Project objectives have strong assessment elements. Improving student achievement is one of those, and it is based on guidelines for outcomes assessment in major programs. After meeting with the president to discuss assessment, attending workshops, and

Assessment Update
Winter 1989
Volume 1, Number 4

consulting with nationally recognized experts, half the faculty (representing 40 programs) had developed assessment plans by early 1989. Students have also been included in workshops and involved in assessments of major academic programs. Students are informed that the assessment process will be interesting, that it will not affect them academically, and that the information will be used to strengthen program content and teaching methods. Assessment methods involve writing, oral communication, and small-group discussion. A comprehensive survey of faculty, students, and alumni will be repeated every five years.

The second of the three is the objective to improve recruitment and persistence of minority students. It is achieved in part through monitoring the success of admissions outreach and a bridge program for urban high school juniors. A system called SAFE (Student Academic Feedback/Evaluation) has been created to provide students and faculty with information that will help improve students' classroom performance. As a result of the Excellence and Equity Project, the number of black and Hispanic students admitted, as well as their persistence toward graduation, has increased. A range of assessment initiatives is planned, to ensure that these gains are maintained and improved.

As a result of Kean College's Excellence and Equity Project, the number of black and Hispanic students admitted, as well as their persistence toward graduation, has increased.

Enhancing the General Education Program is the third objective with a strong assessment component. Student assessment measures have been created by the faculty for both the content and the critical-thinking skills associated with each core curriculum area. Essay tests evaluate five areas: composition, emergence of the modern world, intellectual and cultural traditions, inquiry and research, and landmarks of world literature. A multiple-choice test was developed in the area called "Science and Technology in the Modern World." Initial results of these assessments have been used to evaluate course objectives, analyze curriculum materials, discuss teaching methods, and clarify relationships among courses. The next step is to use the results to assess the program as a whole, refine the assessment process itself, and inform students about their performance. Faculty and student surveys regarding attitudes toward the General Education Program will also provide information.

The Kean College project is clearly a comprehensive effort in organizational development. Its success to date provides an excellent model for change in higher education and for the appropriate role of assessment in this process. For more information, contact Dr. Donald Lumsden and Dr. Michael Knight, coordinators, Assessment of Student Learning and Development, Kean College of New Jersey, Union, NJ 07083. Tel: (201) 527-2661.

Assessment Update
Winter 1989
Volume 1, Number 4

PETER J. GRAY

James Madison University

The comprehensive assessment program at James Madison University (JMU) was developed not only to meet Virginia's state assessment guidelines but also to ensure that JMU's mission is carried out. The program is particularly intended to demonstrate accountability for the proper and wise use of university resources.

The comprehensive assessment program at James Madison University is particularly intended to demonstrate accountability for the proper and wise use of university resources.

JMU has an enrollment of over 10,000 on-campus students, and nearly 80 percent are from Virginia. Predominantly an undergraduate institution, JMU offers a diverse general studies program in the humanities, sciences, education, psychology, business, communication, the arts, health, and human development.

The Office of Student Assessment was created to coordinate campuswide assessment regarding the identification of high-risk students, the assessment of specific academic and student affairs programs, and the evaluation of the campuswide undergraduate programs in general education. Faculty in these programs are instrumental in developing and implementing the assessments.

The identification of high-risk students is intended to create a model that predicts academic success and persistence. The model is based on students' personal and family characteristics, their metacognitive development, and other, more traditional variables, such as high school rank.

The first step in the assessment of academic programs is for faculty to define specific objectives for student outcomes that the program intends to foster. In this process, it is not necessary that every faculty member agree with every objective, nor need all objectives be viewed as equally important. Instead, it is most important for faculty to reach consensus on the general categories of objectives and for individuals to be willing to accept particular objectives, even if they do not fully endorse them.

Objectives cover a broad range of educational outcomes, including skills, knowledge, personal and social development, and ability. Objectives are first

Assessment Update
Spring 1990
Volume 2, Number 1

BEST COPY AVAILABLE

defined in a general way, to create a model of the whole curriculum. Then they are refined to describe specific course-related outcomes. The refinement of objectives is followed by the selection and/or development of techniques to measure their accomplishment. Some measures are used each semester; others are used annually (to assess graduating seniors, for example).

JMU faculty have a wealth of experience in developing and selecting assessment measures. Measurement techniques include standardized, norm-referenced, subject-specific achievement tests; national licensure exams; assessments of thinking skills; formal paper reviews; ratings of personal characteristics, development, and life styles; locally designed achievement tests; personal histories; technical competence and proficiency screening; essay examinations; problem-solving tests; exit interviews; faculty ratings of field and clinical work; oral and written communication evaluations; case presentations; and adviser checksheets.

Most programs at JMU are being assessed through alumni and student surveys, external peer consultations and reviews, and student outcomes.

Most programs at JMU are being assessed through alumni and student surveys, external peer consultations and reviews, and student outcomes. Alumni rate their overall experiences at JMU and their satisfaction with their major programs of study. They indicate their employment and continuing education activities. They also identify the extent of their satisfaction with such general education-related outcomes as effective writing and speaking, use of mathematics in everyday life, problem solving and creative thinking, maintenance of physical health, use of leadership skills, and appreciation of diverse points of view.

A rating form is used to evaluate departmental assessment activities. As criteria, this form lists clarity and comprehensiveness of objectives, comprehensiveness of assessment information, reliability and validity of assessment methods, nature of assessment results, and use of results.

An assessment of student affairs programs has also been initiated. The focus of this assessment is students' development with respect to programs in admissions, residence life, student activities, minority support services, and career planning and placement. The first step is to inventory the student activities and areas of development for which each student affairs office is responsible.

The program in general education is intended to provide a liberal studies foundation for students' undergraduate education. Five developmental areas define the categories of outcomes for this program: identity, or one's sense of self; intellectual development; psychological maturation; moral development; and critical thinking. Students' progress in each of these areas is assessed with measures developed at JMU. For example, identity is assessed by use of an instrument with three subscales—confidence, sexual identity, and conception about body and appearance. Intellectual development is assessed with a forced-choice instrument based on Perry's four stages—dualism, relativism, commitment, and empathy. Psychological maturation is assessed in terms of students' ability to establish and clarify purposes, develop mature interpersonal relationships, realize academic autonomy, attain intimacy, and

Assessment Update
Spring 1990
Volume 2, Number 1

achieve a salubrious lifestyle. In addition to these developmental areas, more traditional areas, such as knowledge in liberal studies courses and writing and computing skills across the curriculum, are also assessed.

JMU has begun to create a very comprehensive assessment of its undergraduate programs in general education, academic disciplines, and student affairs. For more information about these efforts, contact Dary Erwin, Office of Student Assessment, James Madison University, Harrisonburg, VA 22807. Tel: (703) 568-6706.

JMU has begun to create a very comprehensive assessment of its undergraduate programs in general education, academic disciplines, and student affairs.

Assessment Update
Spring 1990
Volume 2, Number 1

The Ohio State University's Center for Teaching Excellence

None of the center's assessment activities focuses on student outcomes as measured by standardized instruments or normative tests.

The Ohio State University's Center for Teaching Excellence conducts a variety of assessment studies related to its mission of providing leadership and services for the enhancement of effective teaching. As with other such centers at major research universities, assessment activities are not limited to student outcomes assessment. In fact, none of the center's assessment activities focuses on student outcomes as measured by standardized instruments or normative tests.

The Ohio State University has over 38,000 undergraduate and 12,000 graduate students enrolled in 19 degree-granting colleges at the Columbus campus. Nearly 98% of the undergraduates are from Ohio. In its current configuration the Center for Teaching Excellence has been in operation for four years, but its history stretches back nearly 50 years. Its four goals are to stimulate and sustain efforts directed at developing and evaluating instructional practice; to introduce and support effective application of instructional technology; to encourage and contribute to the continuing development of faculty and teaching assistants as college teachers; and to initiate projects, conduct inquiries, and disseminate information for enhancing instructional knowledge and practice.

Several different types of assessment data are collected relative to these goals, including student assessment, in the form of cognitive, affective, and skill-related outcomes; retention and other demographics; and student evaluations of instruction. The Center for Teaching Excellence provides assistance through workshops, individual consultation, and written materials in the construction and interpretation of objective tests, essays, performance measures, and other means used to assess student learning in courses and larger programs. Student retention data for the last 15 years have been analyzed to determine general trends and specific differences among various subpopulations. Students also provide assessment information for course improvement in response to "tailored" evaluation instruments concerning instructional materials and procedures.

Assessment Update
Summer 1990
Volume 2, Number 2

– In addition, center staff develop program and facilities evaluation instruments and assist in the interpretation and utilization of results. Programs assessed include library user-education, foreign language skills, clinical problem solving, science concepts and applications, teaching for black student retention, and physical education demonstration. Facilities assessments include lecture halls, standard classrooms, high technology classrooms, seminar rooms, laboratories, and studios.

In addition to the program and facility assessments in which students participate, faculty are involved in peer observations and instructional materials reviews. They may also take advantage of self-assessment services offered by the center, such as videotaping and in-class observation. As teaching assistant supervisors, some faculty provide input regarding the quality of teaching assistant supervision and ways that it can be improved.

In addition to the program and facility assessments in which students participate, faculty are involved in peer observations and instructional materials reviews.

As part of an accreditation self-study during 1985 and 1986, the Ohio State University used the College Student Experiences Questionnaire (CSEQ) to provide a representative picture of the undergraduate population. This instrument was used along with other data that the University routinely collects to describe its students, including data on mathematics and English placement test scores, ACT and SAT scores, course enrollment profiles, enrollments in majors, grade point averages, and graduation profiles. The CSEQ was selected because it provides, in addition to student demographics, student perceptions of institutional climate, student descriptions of the quality of their educational experiences, and student judgments of educational gains at a specific point in time.

In January 1986 the CSEQ was mailed to a stratified random sample of all undergraduates on the Columbus campus. A 33% return rate was achieved and the profile of the respondents was found to closely match the total undergraduate population in terms of class status, major, gender, and ethnic group. The results were shared with university-wide curriculum committees, the college deans, and a student retention committee appointed by the provost, and helped to stimulate discussion on campus. Topics included the quality of education that students experience; the role of curriculum and instruction in meeting the students' general education needs; the institutional climate for both students and faculty; the assessment of teaching and learning; strategies for achieving quality education; and the importance of student effort, both in and out of class, in the educative process. These CSEQ data will provide the baseline for a longitudinal student assessment study. As such, the CSEQ fits well with the other formative assessments of the center.

For more information about the center's assessment and other activities, contact G. Roger Sell, Center for Teaching Excellence, The Ohio State University, 15 Lord Hall, 124 West 17th Ave., Columbus, OH 43210-1316. Tel: (614) 292-1316.

Assessment Update
Summer 1990
Volume 2, Number 2

The Evergreen State College Olympia, Washington

In 1989, Washington's Higher Education Coordinating Board concluded that standardized tests do not contribute to the improvement of teaching.

Assessment Update
Fall 1990
Volume 2, Number 3

The Evergreen State College in Olympia, Washington, is a public baccalaureate liberal arts institution that enrolls approximately 3,000 students. For the last three and one-half years, faculty, students, and administrators in the Assessment Study Group, led by the Evergreen Office of Institutional Research, have been focusing on the central objective of evaluating teaching and learning. This focus had two initial influences: Evergreen's own strategic plan, and the legislative mandate to the Higher Education Coordinating (HEC) Board that Washington State higher education be more accountable to the citizens of the state.

In 1989, HEC concluded that standardized tests do not contribute to the improvement of teaching and will not be required at any level. In addition, HEC chose academic improvement as the only goal of assessment, rather than making it a dual goal, along with public accountability. Nevertheless, institutions will ultimately be held accountable for providing evidence that higher education makes a significant difference to students, but such accountability will be in the context of institution-specific assessment that contributes to the improvement of teaching and learning.

Evergreen strives to produce graduates distinguished by their ability to communicate, conceptualize, and solve problems; by their comfort with diversity and complexity; by their self-reliance as learners and researchers; and by their commitment to personal integrity and the public good. This process begins with the general education learning objectives concentrated in the core programs. These introduce first- and second-year students to the central mode of study at Evergreen—seminars in which faculty from three of four different disciplines (humanities, arts, social sciences, natural sciences) explore a central theme or problem in a coordinated way.

This mode of study continues through interdisciplinary learning communities, which immerse students in complexity and in diversity of perspectives and are intended to foster development of cooperation, communication, and integration. Small classes, independent study, internships, and applied projects are used to bridge theory and practice. Narrative grading and self-evaluations require active involvement of students from the start in the assessment of their own learning. Several principles guide the work of the Assessment Study Group. First, to be valuable, assessment must be useful to the institution, especially to faculty and students. Second, it should be based on multiple measures. Third, it should be predominantly if not exclusively "homegrown." Fourth, it should emphasize qualitative databases.

Assessment projects reflect the Evergreen learning environment and the guiding assessment principles. For example, a videotaped documentary has been produced in which faculty, students, and alumni discuss the nature of teaching and learning at Evergreen. Surveys of current students and alumni have also been completed. Some of the questions on the alumni survey, for example, correspond to items contained on the American College Testing Alumni Survey. Therefore, national norms are available for purposes of comparison. The response profiles indicate ways that Evergreen students are alike and ways that they are different from other public college students.

To be valuable, assessment must be useful to the institution, especially to faculty and students.

Content analyses of Evergreen's narrative transcripts are being conducted, to identify and evaluate the common elements of the teaching and learning environment. In another study, the Perry scheme of cognitive and ethical development is providing the conceptual model for the analysis of students' self-evaluations, which are completed at the end of each program. This study is attempting to show how much and in what ways students grow during their years at Evergreen, as well as how their cognitive development may be different from that of students at other institutions.

Still to be implemented are such assessments as a videotaped documentary of teaching and learning over an entire academic year, an ethnographic study that will describe and evaluate the effects of the culture of teaching and learning at Evergreen, a longitudinal portfolio analysis of work that a student completed during her career at Evergreen, a study of group-process skills, and evaluations of writing and quantitative skills using both innovative and traditional methods.

Since these assessments are intended to improve teaching and learning, they often provide practical information that can be used immediately. For example, the study of cognitive development is providing faculty with insights and guidelines for interacting with students within a single seminar who are at various cognitive stages.

For more information about the work of the Evergreen Assessment Study Group, contact Steve Hunter, Office of Institutional Research, The Evergreen State College, Olympia, WA 98505. Tel: (206) 866-6000 ext. 6567.

Assessment Update
Fall 1990
Volume 2, Number 3

The Office of Educational Development at the University of California, Berkeley

A special need at Berkeley is to assess the impact of increased diversity among undergraduates.

The Office of Educational Development (OED) is the primary instructional support agency at the University of California, Berkeley. This unit, established within the Office of Undergraduate Affairs for the last ten years, coordinates the assessment and instructional improvement activities conducted by faculty and administrative units.

A special need at Berkeley, as well as at other institutions, is to assess the impact of increased diversity among undergraduates. Informal and campus-specific assessment and evaluation techniques were used to determine ethnic, cultural, and racial attitudes among students. Small group interviews were conducted with over 230 undergraduates, using a "snowball" sample technique, in which earlier interviewees recommended those to be interviewed later. Results of the interim study report will be used to make decisions about programs and activities that will promote diversity and facilitate students' adjustment to Berkeley. For example it is recommended that students form small groups that meet every three weeks to address problems of orientation, adjustment, and integration into campus life, and that faculty roles in encouraging interaction among students be increased through curriculum-related group projects or study groups.

This year, the Academic Senate and the OED will sponsor a grant program for faculty, to encourage their efforts in assessing lower-division education. Proposals may focus on measuring cumulative learning of Berkeley's undergraduates, determining the quality of contact between instructors and lower-division students, assessing students' achievement and performance during their first two years, examining the coherence of lower-division service and introductory course sequences, evaluating the effectiveness of lower-division courses in preparing students for majors, and studying the concerns and needs of faculty who teach lower-division courses.

Multiple measures and existing data are used at Berkeley to enhance understanding of such complex issues as lower-division education. For example, a major retention study conducted by the Office of Undergraduate Affairs

Assessment Update
Winter 1990
Volume 2, Number 4

used existing data to gain a fuller understanding of students' similarities and differences and the relationship of these characteristics to student retention. Methods included statistical analyses of the effects of race, gender, ethnicity, parental income, and SAT scores on grade point averages and graduation rates. This study also took account of responses to qualitative and quantitative survey items on students' attitudes about education in general and their experience at Berkeley. The results of this study have been widely distributed, both on and off campus, to inform the community about Berkeley's efforts to recruit, admit, and retain students, particularly underrepresented minorities.

Assessment and evaluation are also ongoing and integral parts of many campus programs. For example, diagnostic tests in writing and mathematics are used to place students in appropriate courses. The Office of Student Research routinely conducts surveys of entering freshmen and exiting seniors. The College of Letters and Science also schedules interviews with students who choose to leave the university. In response to a concern about student advising, the Office of Undergraduate Affairs conducted a major review of previous advising studies, dating back to 1972. In addition, in the spring of 1990, the Commission on a Changing Student Body administered a survey to over 1,600 students to assess their reactions to advising support services and staff. This year, the Campus Retention Council will use these results to develop ways to strengthen advising, especially for students who have not declared majors.

Several units, in addition to OED, have ongoing formative evaluation programs to examine the impact of support services. In some cases, these evaluations have led to restructuring of the services themselves, as well as of how they are delivered. In addition, peer reviews of all undergraduate and graduate programs of study are conducted on a regular cycle. Peer reviews of the quality of faculty members' teaching are part of the regular merit-and-promotion process. Surveys of students in every course each semester are used by departments to help determine the effectiveness of instructional materials and the teaching performance of instructors. Of course, as on almost all other campuses, external assessments are regularly conducted by state, regional, and professional accrediting agencies. These assessments often have strong self-evaluation components, which act as a stimulus to local discussion of instructional and support program quality.

Of special importance are assessment studies conducted in response to questions posed by various groups and individuals at Berkeley.

Of special importance are assessment studies conducted in response to questions posed by various groups and individuals at Berkeley, such as the Admissions Office, the Graduate Division, the Office of Student Research, Academic Senate Committees, individual department heads, and senior academic administrators. For example, campus-wide surveys of entering freshmen are used to assess their backgrounds, interests, aspirations, and attitudes, and students' records are reviewed in determining the ethnic diversity of applicants, enrollees, graduates, and students who have dropped out. Such information is basic to understanding the student population and to developing appropriate instructional programs and support services.

The Office of Career Planning and Placement and the Office of Student Research recently surveyed alumni on the quality of education provided at Berkeley. Alumni were asked about their graduate education and career experiences, the aspects of their Berkeley education that were most important to them, and the experiences or courses they wished they had had. They said that basic written and oral communication skills were most critical in their current work or graduate study. The information gathered from this survey was given to individual departments for their use in course and curriculum revision.

To encourage the use of assessment findings in instructional improvement, OED offers individual consultation with faculty about strengthening their teaching skills, developing early feedback methods for their courses, and creating new curricula and course syllabi. In the Topics on Teaching Series organized by OED, discussions led by outstanding faculty have focused on teaching the large lecture class, responding to diversity in the classroom, designing writing assignments and exams, and grading coursework. Smaller, hands-on workshops, led by OED and other staff, have covered such topics as preparing midterm exams, using instructional media, and other skills.

Alumni said that basic written and oral communication skills were most critical in their current work or graduate study.

OED also produces publications on teaching and learning. Large and small grant programs for instructional improvement provide funds for revising teaching strategies, redesigning instructional materials, enriching existing courses, and creating new course procedures and materials. Some grant programs are intended to address specific areas, such as campus diversity, writing and speech, and collaborative learning; others are intended for particular audiences, such as faculty who have received the Distinguished Teaching Award. There is also a travel grant that supports instructors' participation in professional activities involving training in the use of new teaching techniques, resources, or devices.

For more information regarding assessment efforts and related support services at Berkeley, contact Barbara Gross Davis, Dean of Educational Development, University of California at Berkeley, 273 Stephens Hall, Berkeley, CA 94720. Tel: (415) 642-6392.

Assessment Update
Winter 1990
Volume 2, Number 4

PETER J. GRAY

Northeast Missouri State University

Since the 1973-74 academic year, Northeast Missouri State University (NMSU) has utilized achievement and attitudinal instruments to measure student performance outcomes. This assessment program was motivated by the change in the university's mission that occurred in the early 1970s. At that time, NMSU was transformed from a teachers' college to a small multipurpose university. The university is now a selective, primarily undergraduate institution serving mostly traditional students in a residential setting, with a total enrollment of 6,500.

Consistent with its new multipurpose mission was the need to create a system to establish and maintain educational quality in diverse areas of study. In essence, the institution had to identify, develop, and nurture the human and physical resources that would produce nationally competitive graduates. In response to this challenge, the faculty and staff of NMSU designed an outcomes assessment program that focused on student learning.

Three goals guided the development of the program: to know everything possible about students, to demonstrate that the university made a positive difference in students' lives, and to determine whether graduates were nationally competitive. More specifically, the program was intended to develop a model of assessment that would monitor student growth and development beyond course grades, keep the disciplines current, ensure the integrity of the degrees awarded, focus on quality rather than quantity as a measure of institutional success, and build a pervasive campus climate emphasizing student learning.

Faculty identified four important points in a student's university career as targets of outcome assessment activity: upon matriculation, to measure entering achievement and attitudes; upon completion of the sophomore year, to measure growth in general knowledge and attitudes; upon graduation, to measure knowledge, skills, and attitudes derived from the total undergraduate experience; and after leaving NMSU, to gather information from alumni on their satisfaction with their preparation for adult living.

Three goals guided the development of the program: to know everything possible about students, to demonstrate that the university made a positive difference in students' lives, and to determine whether graduates were nationally competitive.

Assessment Update
March-April 1991
Volume 3, Number 2

BEST COPY AVAILABLE

In addition to basic demographic data (age, gender, high school performance, college preparation) used in the analysis of various other kinds of data, three parts of NMSU's student outcomes assessment program have been in place since the mid 1970s. First is the value-added component used to measure student growth in general knowledge from the beginning of the freshman year to the end of the sophomore year. The Collegiate Assessment of Academic Proficiency (ACT/CAAP) and the College Outcome Measures Program (ACT/COMP) exams are administered to students in matched sequence for this purpose. Second is a standardized, nationally normed examination (GRE, NTE, MFAT, or UAP) that provides the comparative component for demonstrating student achievement in the major field. When such an exam is not available, a locally normed senior examination is used. Third is the attitudinal component, which involves determining students' self-perceptions of growth and their evaluations of the university and its services. Several locally developed surveys are used to gather attitudinal information, including the Summer Orientation Student Survey, the annual Institutional Student Survey, the Graduation Student Questionnaire, and three-year follow-up surveys of alumni and employers. These multiple assessment measures are combined with classroom outcomes and observations to provide data for decisionmaking.

Fears were expressed that assessment would become a tool for evaluating faculty, and that faculty would begin to teach to the tests.

Results of the various assessment components are made available to individual student, program disciplines (departments), divisions (schools), and the university. Comparisons can be made longitudinally, over time; among levels; with the university as a whole; and, where data are available, with national norms.

In the process of establishing and developing a comprehensive, university-wide outcomes assessment program, NMSU moved through four stages: the readiness stage (1972 to 1974), the implementation stage (1974 to 1981), the acceptance stage (1982 to 1986), and the commitment stage (1987 to the present). Each stage has had its costs and benefits.

The readiness stage at NMSU included intense debates about what could and should be measured, what should or should not be recorded (to safeguard the use of data), and when and by whom interpretation of data was appropriate. Fears were expressed that assessment would become a tool for evaluating faculty, and that faculty would begin to teach to the tests. Clearly, the doubt, insecurity, fear, misunderstanding, and disagreement that developed, as well as the time spent in meetings and the money spent for consultants, were major costs of the readiness stage. However, the institution also benefited during the readiness stage, through the stimulating discussion of campus-wide goals and objectives. Nevertheless, if the costs and benefits of this stage could be weighed against each other, the intangible costs would seem to exceed the benefits. Administrative support and steadfastness were essential for progress to the implementation stage.

Assessment Update
March-April 1991
Volume 3, Number 2

During the implementation stage, when data were first collected, analyzed, and interpreted, benefits began to accrue for specific programs and for students. However, the benefits were still not sufficient to outweigh the emotional and time costs perceived by faculty. The only way to maintain momentum was by continued administrative support and concrete evidence that the purpose of assessment was to enhance student learning through program improvement, not to evaluate faculty.

The benefits to the university at the acceptance stage were numerous. For example, assessment began to be integrated throughout the university. Faculty began to use the data collected to identify areas in need of improvement. Students increased their expectations, their involvement in learning, and their attention to educational goals. Emotional costs diminished at this stage, even though some faculty and students were still concerned that results did not match goals. Time costs shifted, from time spent on developing assessment to time spent using the information collected for curricular and other revisions. The true benefits and costs of maintaining the assessment program began to become clear at the acceptance stage.

At the commitment stage, assessment is no longer being considered a discrete function; it has become an integral part of nearly every activity of the university. The costs have by now become part of the normal operations of the institution, in terms of student, faculty, and staff activity and institutional financial commitments. Numerous benefits accrue to the institution.

The assessment process and the discussions that it stimulates seem to facilitate active student learning and involvement, feedback on individual student and faculty strengths and weaknesses, cooperative learning among students and faculty, closer integration of the curriculum, and professional and scholarly interaction among faculty and students. In addition, discussions focusing on student outcomes have changed the ways in which the university serves and involves students in campus living and learning experiences. Administrators, faculty members, and students utilize assessment resources in planning, curriculum revision, development of research grant proposals, and the whole set of activities connected with learning and living at the university.

Discussions focusing on student outcomes have changed the ways in which the university serves and involves students in campus living and learning experiences.

The review of longitudinal data has revealed trends related to the mission of the university and the future direction of undergraduate education at NMSU. After almost 20 years of assessment, NMSU has moved into a new era, with programmatic reconfiguration and a mission change intended to establish NMSU as the selective public liberal arts and sciences institution of Missouri. Today the evaluation of both the curriculum and the ambience of the university focuses on the liberal arts and what they mean for a university.

While many elements of the original program have been retained, the focus of the current assessment effort has shifted considerably. The focus has moved from university-conducted outcomes assessment to involvement of students in self-assessments of individual academic progress. The student portfolio is a major element used to accomplish this shift. Portfolio assessment brings the process full circle, to the critical issue of student outcomes, but in a context where the teacher and the student become collaborators in achieving a common objective-student learning.

For more information, contact W. Jack Magruder, Vice-President for Academic Affairs, Northeast Missouri State University, Kirksville, MO 63501. Tel: (816) 785-4105.

Assessment Update
March-April 1991
Volume 3, Number 2

Miami-Dade Community College

A series of reforms led to the current assessment system.

Miami-Dade Community College was founded in 1960 as an extension of the public school system and had an opening enrollment of 2,025 students. In 1968, Miami-Dade separated from the public school system and has been governed since then by a board of trustees as part of the community college system in Florida. In 1985, Miami-Dade graduated its 100,000th student, enrolled its 500,000th credit student, and celebrated its 25th anniversary. The credit student headcount on the five Miami-Dade campuses grew to 71,990 for the 1988-89 academic year, and the noncredit headcount was 45,948, for a total of 117,938 students.

A series of reforms led to the current assessment system, beginning with the General Education Study Committee, formed in 1975 to address recommendations from the 1974 Institutional Self-Study for a thorough review of the general education program. The general education program was designed to serve diverse community needs in the context of an open-door admissions policy.

In order to involve as many faculty as possible in establishing goals with regard to curriculum changes, each campus established general education committees. On South Campus, for example, every faculty member was assigned to one of nine subcommittees. By 1978, the general education proposal, which addressed new basic skills requirements and embodies a new general education program, was endorsed by the College Committee on Academic Affairs, the president and president's council, and the Miami-Dade District board of trustees.

The new curriculum was implemented in 1981. This curriculum separated the general education credits into core requirements, distribution requirements, and electives. New general education courses were developed by faculty committees.

While Miami-Dade has always used an assessment instrument for course placement with new students, the new curriculum re-emphasized the

Assessment Update
May-June 1991
Volume 3, Number 3

importance of basic skills assessment at the time of admission, for the purpose of determining a student's competence in reading, English, and mathematics. Developmental courses were designed to help students remedy deficiencies in any of these areas.

The assessment tool had always been used for advising students about which courses were appropriate for their academic programs. Before 1981, however, students were not required to follow this advice. After 1980, that policy changed, and placement into necessary college preparatory courses became mandatory at Miami-Dade. At the same time, a student flow model was developed to help students understand how they could move successfully through the curriculum.

Success in implementing these new requirements depended on the development of related support systems. At Miami-Dade, three assessment systems were developed to support the student flow model, including the academic alert system, the advisement and graduation system (AGIS), and the standards of academic progress (SOAP) system.

Academic alert, a midterm progress report sent to every enrolled student, is generated from instructors' feedback on students' attendance and academic progress. If progress is reported to be unsatisfactory, the personalized letter suggests sources of assistance. Because students are informed of their progress at midterm, they still have ample time to obtain the assistance they need to be successful in their current courses.

At Miami-Dade, three assessment systems were developed to support the student flow model.

AGIS is an online report of degree requirements, based on the student's degree and transfer intentions. The report combines information from the student's master record and transcript files, to align graduation requirements and status by requirement area, rather than by transcript-like chronological order. Another component of the AGIS report informs students of the transfer requirements for related majors at four-year institutions. A course sequencing pathways document, the third component of AGIS, identifies, by area, courses that are required or recommended, on the basis of assessment results, the student's major, and the institution to which the student plans to transfer.

At the end of each term, the SOAP system is used to review students' progress in course completion and academic achievement. Students who are not progressing or achieving satisfactorily are required to see an adviser and are restricted in their enrollment or separated from the college for a specified period of time.

During the past decade, the Florida state legislature has become involved in establishing entrance and exit requirements for students in higher education. Miami-Dade chose to use the Multiple Assessment Programs and Services (MAPS) test as the assessment tool and continued to require mandatory course placement on the basis of results. Such placement is now also mandated by the legislature. In 1986, Miami-Dade developed a partnership with IBM and the College Board to pilot a program for assessing students' basic skills via computer. As a result, Miami-Dade now uses the Computerized Placement Test (CPT) as its assessment instrument, having received permission from the State of Florida to do so.

The State of Florida requires additional testing for students who wish to earn the associate of arts degree or who wish to become juniors at a state university. The College Level Academic Skills Test (CLAST) measures achievement in computation, communication skills, reading, and writing. The first administration of the CLAST took place in October 1982. The test was used to establish mean scores for the multiple-choice subtests and the essay score. Beginning in October 1984, the student had to meet established passing standards in order to graduate with an associate of arts degree or become a junior at a state university. Minimum score requirements were raised in August 1986 and were due to be raised again in August 1989.

A key issue is the effect that raising standards will have on the graduation rate at Miami-Dade and other campuses. It is safe to assume that this rate will decrease dramatically, especially for minority students, if higher standards are adopted. In fact, research by the Miami-Dade Office of Institutional Research indicates that there is a strong inverse relationship between CLAST scores and minority group affiliation. Analyses have shown that minority students score lower on the test, that raising standards has a disproportionate negative effect on minorities, and that the CLAST does not predict minority performance in the state university system as well as it does that of other students. This study was conducted with only those students who had completed at least 60 credits and who therefore could be judged as successful on the basis of grades and course completion.

A key issue is the effect that raising standards will have on the graduation rate at Miami-Dade.

At the heart of the debate over CLAST is the issue of predictive validity—that is, the relationship between scores and the students' success. Research conducted by the Florida Inter-Institutional Council on students who satisfied the 1986 standards but not the 1990 standards indicates that 78% of the 1,771 students studied had a GPA of C or higher during the first term at a state university.

Two new collegewide projects, the Teaching/Learning project and the New Directions project, have been defined by Dr. Robert McCabe, president of Miami-Dade, as "Reform II." The Teaching/Learning project, with the involvement of faculty from each campus, seeks to improve the quality of teaching and learning at the college, to revitalize and emphasize teaching as a rewarding career, and to establish teaching at the center of college activities. Toward this end, the Endowed Teaching Chairs program, among the first in the community college system nationally, has been implemented to reward faculty for excellence.

The New Directions project, also involving faculty from each campus, has begun to look at assessment and curriculum issues. Committees have been formed to address needs of students and faculty in such areas as the honors curriculum, college-preparatory disciplines, CLAST preparation, the content of core courses, and student flow.

Assessment Update
May-June 1991
Volume 3, Number 3

The faculty partnership with administrators in creating new programs and policies has meant that Miami-Dade has been able to make massive changes in its collegewide curriculum, streamline support services to provide better information to students, reemphasize the classroom as central to the college's mission, and provide more effective promotion and evaluation. For more information, contact Karen Hays, associate dean for career planning and advisement, Miami-Dade Community College, South Campus, Miami, FL. Tel: (305) 347-2134.

The faculty partnership with administrators in creating new programs and policies has meant that Miami-Dade has been able to make massive changes.

Assessment Update
May-June 1991
Volume 3, Number 3

Ohio University's Multidimensional Institutional Impact Assessment Program

The University's assessment of quality must have consequences and must offer direction for decisionmaking and action, in order to be valuable to the University.

Assessment Update
September-October 1991
Volume 3, Number 5

Guest Columnist: A. Michael Williford, director, Office of Institutional Research, Ohio University.

Over the last ten years, Ohio University has developed a unique approach to student assessment. Currently a leader in the assessment movement in Ohio because of its early commitment to centralized student assessment and the flexibility of its program, Ohio University is the oldest public university in Ohio. Enrolling over 26,000 students at its main and regional campuses, the university offers a diverse academic program with over 250 academic majors and a strong emphasis on liberal education. The main campus is residential, located in southeast Ohio, serving about 15,000 undergraduate and 3,000 graduate students. Eighty-five percent of the students are from Ohio.

The state of Ohio does not mandate assessment; Ohio University's commitment to assess its impact on its students began as a result of the 1980 State of the University address delivered by President Charles J. Ping to faculty, students, and staff. The president called for the University to examine its commitment to growth in quality, stating that the University's assessment of quality must have consequences and must offer direction for decisionmaking and action, in order to be valuable to the University.

President Ping appointed a task force of faculty and staff to spend one year developing a program for assessing the impact of the college experience on students at Ohio University. This program addressed specific goals in the University's ten-year educational plan. A multidimensional program called the Institutional Impact Project, conducted by the Office of Institutional Research, was proposed, accepted, and funded beginning Fall Quarter 1981.

Institutional Impact Project Components

Ohio University uses the ACT College Outcome Measures Program (COMP) instrument to assess Ohio University's general education program. The

BEST COPY AVAILABLE

University began using the COMP Objective test in 1981, testing random samples of freshmen and seniors in a cross-sectional design until 1984, when longitudinal testing of seniors was begun. Over the last seven years, Ohio University has retested seniors who took the COMP as freshmen.

The Student Involvement Questionnaire is used to measure student involvement and integration at Ohio University. This questionnaire collects information in three areas: academic involvement, social involvement and activities, and students' commitment to and satisfaction with the institution.

Analysis of retention and attrition patterns is essential to Ohio University's assessment program. The Student Tracking System utilizes a longitudinal cohort survival method of tracking all freshmen until they either graduate or drop out of the University. Its purpose is to identify potential areas of high attrition and to monitor progress with the University's retention programs from year to year. Retention of freshmen at Ohio University has improved steadily since 1977.

Analysis of retention and attrition patterns is essential to Ohio University's assessment program.

The Freshman Treatment Study collects information from new freshmen about perceptions of treatment received from other people, quality of information they receive, and processes they go through while at Ohio University. It is administered every three years, and changes in scores are plotted from one freshman class to another.

Two separate alumni studies were developed. The first was designed to collect information on graduates' employment status, graduate or professional school acceptance, and employment satisfaction, within one year of graduation. A more detailed Alumni Survey was developed for graduates after they had been away from the University for at least five years. This survey asks questions about traditional alumni outcomes, such as type of job held, employment status, salary, employment satisfaction, graduates' programs of study, and their current satisfaction with their undergraduate programs. The survey also contains questions specific to each academic college's programs; eight separate questionnaires exist, one for each college.

Use of Findings

COMP results provided some useful evaluative information about general education during a change in the general education curriculum. The general education program underwent a review that was implemented in 1979, and the COMP was introduced at Ohio University in 1981. An increase in COMP scores coincides with the implementation of the new general education requirements program: seniors who took the COMP before the new general education requirements scored lower than students who took the COMP after the new general education requirements were implemented. This study was used in the University's application for Ohio "Program Excellence" funding to further enhance its general education program, which it received in 1990.

Assessment Update
September-October 1991
Volume 3, Number 5

Involvement data have been used in three ways. First, a year-to-year comparison of freshmen has been conducted. Second, freshman-to-senior longitudinal comparisons have been made: seniors who participated as freshmen are asked to participate again. Third, the information is used as part of an effort to increase retention. Individual freshman responses are used to identify students who are in good academic standing each year who might voluntarily leave, so that student affairs staff and advisers can intervene and help some solve attrition-related personal problems.

Involvement, attrition, and treatment assessment information is used by Ohio University's academic colleges and departments to develop their own retention programs. For example, information from the Treatment Study was used to establish a new advising program for undecided students in the University College, while the Student Tracking System gives each college an annual progress report on the effectiveness of its retention programs. Information from all these assessment studies has contributed to improved programs and retention over the last ten years; freshman retention was over 85% in 1990.

Information from assessment studies has contributed to improved programs and retention over the last ten years; freshman retention was over 85% in 1990.

Academic college- and department-specific placement information is shared with academic deans and department chairs, who use it for advising, accreditation reviews, academic planning, and reviews by the University Curriculum Council. The admissions office and individual colleges use this information for advising prospective and current students, respectively. The Alumni Survey is being used to evaluate the success of academic programs on a long-term basis and to identify areas of competence that are important to Ohio University graduates several years after they graduate.

All major planning units on campus receive a variety of reports on the Institutional Impact Project. Each academic unit receives personalized reports on its own students and graduates, which are related to specific planning needs. The University Planning Advisory Council receives periodic briefings on the Institutional Impact Project in order to make better resource allocation decisions. Trustees, the Faculty Senate, the University Curriculum Council, the National Alumni Board, and the University Advising Council also receive reports. Individual campus offices, departments, and committees receive information produced from the assessment system. Each office uses results uniquely for planning and evaluation purposes.

Ohio University's Institutional Impact information has been instrumental in its receipt of recent Program Excellence awards from the Ohio Board of Regents. Between 1985 and 1989 Ohio University won four of these awards, totaling about $650,000, for four "excellent" academic programs. In 1990 Ohio University was awarded over one million dollars in an Eminent Scholar award and four Program Excellence awards.

Assessment Update
September-October 1991
Volume 3, Number 5

Future Directions

Ohio University has gained statewide recognition for its Institutional Impact Project. In March 1988 the University sponsored a conference on student outcomes assessment that was attended by most of Ohio's two- and four-year colleges and universities. Since the Ohio Board of Regents first turned to Ohio University in establishing an assessment task force for the state, examples from the Institutional Impact Project and their use at Ohio University will shape how assessment is conducted in Ohio.

The Institutional Impact Project is evaluated regularly, and instruments are added and modified in areas in which information needs change. For example, the need for discipline-specific assessment information is apparent, and ways of gathering this information are being sought. In addition, an assessment program is being developed for each of the regional campuses.

Ohio University's Institutional Impact information has been instrumental in its receipt of recent Program Excellence awards.

Ball State University

Assessment at Ball State focuses on the general studies program, which enhances the potential for effective assessment.

Assessment Update
November-December 1991
Volume 3, Number 6

Ball State University is a public, residential institution of nearly 18,000 undergraduate and 2,000 graduate students. The Office of Academic Assessment was funded by the Indiana legislature in 1987 through a special "excellence in education" allocation. The Office of Academic Assessment, which employs four full-time assessment professionals, and a parallel Office of Institutional Research report to the assistant provost.

General Studies Assessment

Assessment at Ball State focuses on the general studies program, which enhances the potential for effective assessment. Characteristics of the program include a consistent program among all six colleges; specifically defined goals and objectives; a core curriculum program, with five core courses and very limited distribution options; and a junior-level writing competency examination for all students.

The Undergraduate Education Committee is responsible for implementing, monitoring, and evaluating the general studies program, which involves both student testing and program evaluation. Student assessment methods include the College Basic Academic Subjects Examination (C-BASE), a standardized achievement test, and the National Center for Research to Improve Postsecondary Teaching and Learning (NCRIPTAL) Student Goals Exploration (SGE), which are used at the time of entrance and at the junior level in a pre-/post-test design. Both instruments are closely aligned with the general studies goals and objectives.

The C-BASE is administered to students at entry, during the summer before their freshman year, and at exit to juniors or seniors. The test yields 13 specific subject or cluster scores and three general competency scores. Cluster scores are consistent with distribution requirement areas, while general competency scores are related to interpretive, strategic, and adaptive reasoning. The C-BASE results illustrate, first, the preparedness of entering students and, second, the relative achievement of students at the end of their general studies program.

The SGE is based on three premises: that the goals students have for attending college and for enrolling in specific courses have an impact on educational outcomes; that these goals can be measured with a comprehensive survey instrument; and that when measures of student goals are available, institutions should be able to relate them to other student characteristics and so guide efforts to improve teaching and learning. This instrument has been used with about 2,000 students enrolled in general studies courses.

A related effort, not strictly tied to assessment, is the Making Achievement Possible (MAP) project, which involves the administration of a computerized survey to incoming freshmen during the first four weeks of classes. The survey asks students to report their expectations about academic performance and to give self-ratings of academic and personal skills, as well as to report their expected personal involvement on campus. The results of the survey are combined with basic admissions data (high school rank and test scores) to produce two reports.

The student receives a personalized document that reports his or her responses, admissions data, and results from a Learning and Study Strategies Inventory (LASSI) taken during summer orientation. Academic advisers and residence hall directors receive a two-page summary of responses. In addition, a report of those data for the entire class is produced for general institutional review. The results of the MAP project also have potential for use in predicting academic success related to general studies and departmental majors and for specific examination of subgroup profiles related to academic success.

The educational outcomes of academic programs are also assessed at the single-course and departmental levels.

Departmental Assessment

The educational outcomes of academic programs are also assessed at the single-course and departmental levels. Each department is expected to develop an assessment plan, with departmental and course goals and objectives, and a timeline of assessment activities. A total of $40,000 has been made available to support the development of assessment plans. Program reviews for departments with general studies courses will occur on a three-year cycle, with one-third of the program departments being reviewed each year, beginning in 1991-92. All departments prepare reviews on a five-year cycle for the Indiana Commission on Higher Education.

Departments are encouraged to use a wide variety of assessment data sources and data collection tools. Sources include existing departmental data in the form of student transcripts, the results of common course examinations or assignments, student patterns of course selection and facilities use, course enrollment statistics, and graduation rates. Existing student data, such as SAT or ACT scores and high school percentile rank or GPA, as well as new data from such sources as major field tests and the National Teacher

Examination, can provide a comparative picture of departmental effectiveness. Results from the C-BASE and the SGE are also available for departmental use. Each departmental assessment plan contains unique combinations of existing sources of data and of data collected by newly developed assessment tools.

Overall Assessment

Several techniques are used for assessing students' overall experience at the university, including surveys of incoming freshmen's and graduating seniors' attitudes, surveys of students' skills and perceptions, and surveys of faculty, alumni, and employers. These groups are also involved in structured face-to-face and telephone interviews.

Assessment techniques also include a junior-level writing competency examination and a comparative analysis of data from standardized tests. Pre- and post-test data from the SGE are also used in program assessment, as are longitudinal profiles developed from the MAP project. Results from these assessments, which focus on overall strengths and weaknesses of a Ball State education and not on instructors or students, are distributed throughout the university.

Results from assessments, which focus on overall strengths and weaknesses of a Ball State education and not on instructors or students, are distributed throughout the university.

The data that have already been collected provide information that has been used in the formative development of the general studies program and in the improvement of departmental courses and majors. By the end of the 1993-94 academic year, all departments will have been reviewed, and other general studies program data will have been collected for several years. At that time, the general studies program goals and objectives will be reevaluated.

For more information about the assessment program at Ball State University, contact Beverley Pitts, Assistant Provost and Executive Director of Research or Catherine Palomba, Director of Academic Assessment and Institutional Research, Ball State University, Muncie, IN 47306-0241. Tel: (317) 285-3716 (Pitts) or (317) 285-1337 (Palomba).

Assessment Update
November-December 1991
Volume 3, Number 6

PETER J. GRAY

State University of New York College at Fredonia

The State University of New York College at Fredonia is a comprehensive undergraduate college of approximately 4,500 students and 230 full-time faculty. Almost all of the students are from New York State, with about 8% representing minority groups.

Assessment at Fredonia has been concerned primarily with the new General College Program adopted in January 1982 by the Faculty Council. This program replaced a typical distribution-based general education program when concerns were raised about the quality of the experience students were getting. Minda Rae Amiran, former dean for Liberal and Continuing Education, was responsible for the initial development of SUNY Fredonia's approach to assessment.

The General College Program (GCP) has three parts. Part 1 includes courses that introduce the ways people think, write, read, and use numbers. Courses in Part 2, the general education component of the program, offer introductions to fields of study and provide the context for further development of the abilities and competencies identified in Part 1. Part 3 courses are at the 300 and 400 level and are taken by majors and nonmajors alike. They focus on enhancing intellectual, analytical, and communicative powers such as the ability to identify and question basic assumptions or to examine and minimize the influence of biases, ethnocentric viewpoints, and stereotypes. In essence, the GCP is not a core curriculum, but three sets of skills that the courses in each part should promote.

The Fund for the Improvement of Postsecondary Education (FIPSE) provided support for measuring the extent to which students have accomplished the sets of skills of the GCP. The basic purpose of the assessment was to determine whether students developed as thinkers and learners beyond the level they had attained as entering freshmen and, if so, whether this development could be attributed in part to their education at Fredonia. The assessment took place during the 1988-89 academic year.

Concerns were raised about the quality of the experience students were getting.

Assessment Update
March-April 1992
Volume 4, Number 2

BEST COPY AVAILABLE

Student development was determined by comparing freshman and upperclass assessment results on paper-and-pencil tests designed by the campus Assessment Committee. During the development of the tests, the Assessment Committee distributed to academic departments definitions and readings about the topics to be assessed, as well as test descriptions.

The tasks comprising these tests were designed to measure students' ability to write, read, analyze their own thinking and learning, solve numerical problems, reason scientifically, and demonstrate socioethical understanding (two tests). A sample of 40 first-year students and 40 statistically similar upper-division students (juniors and seniors) took each test, for a total of approximately 320 in each area. Scores on the same tests by matched groups of freshmen and upperclassmen in a comparable public university were used to help the Assessment Committee understand the influence of maturation in any college environment on the results.

Given that all of the assessment instruments were paper-and-pencil tests, to some extent the entire assessment program is a writing assessment. However, the criteria used to assess responses to the tasks of each test reflect the relevant goals and objectives of the General College Program.

Student development was determined by comparing freshman and upperclass assessment results on paper-and-pencil tests designed by the campus Assessment Committee.

In all the tests, Fredonia faculty teams of two readers scored the responses blindly, that is, without identification of individual students' identity, class status, or college affiliation. Interrater reliability and the statistical significance of the difference between groups of students on each test and on each part of each test were calculated. Interrater reliability is reported to have been very high for all the tests.

The writing test served as a model for the other tests. The writing test involved an impromptu 50-minute assignment of describing and analyzing what students had found to be a major problem with the educational system in their high school. The criteria used to assess the resulting essays were first a holistic view of the essays' quality and then specific criteria regarding students' ability to formulate a central point, organize the essay in relation to that point, present supporting evidence, analyze reasons, and write correct English.

The tasks and criteria of the other tests are summarized below to show the scope and depth of this locally developed assessment program:

Reading. Students had 1-1/2 hours to read a 12-page, nontechnical professional journal article on social differentiation in high school and to answer questions about it, for example, its main point, the reasons the author gave to establish that point, the implications of the argument, the assumptions, and the organization and style of the essay. Students were asked to indicate whether they agreed or disagreed with the article and were also asked to provide an example from their own experience illustrating or contradicting one of the author's points.

Assessment Update
March-April 1992
Volume 4, Number 2

Reflecting Thinking. An hour-long, two-part task was used to test students' ability to analyze their own thinking and learning, through the metacognitive skills of self-awareness, self-monitoring, application, and evaluation. Part One asked students to describe their experience of learning something outside school, and Part Two asked five questions that led students to review what they had written (for example, What were the major factors influencing their nonscholastic learning experience? What were the differences between learning in the experience they described and in school? What conclusions could they draw about learning from the experience in general?).

Quantitative Problem Solving. This test was different from the others in its focus on mathematics. Answering the test questions involved using simple algebra and estimation. However, scores were based on students' solutions to the questions, explanations of their assumptions, and on the reasoning used in arriving at solutions.

Scientific Reasoning. To test the achievement of this outcome, students had to design an experiment and critique a correlation study report. Assessment criteria for the experiment included the extent to which students varied the conditions, controlled the variables, selected a reasonable sample size, described suitable measures of growth, and gave the experiment a reasonable duration. The critique was assessed in terms of the extent to which students questioned the sample characteristics, research methodology, statistical significance, and research bias. Reviewers also looked for suggestions regarding other factors that could have caused the observed correlation.

Socioethical Understanding. Three sets of questions were used to assess this area. In the *History* set, students were asked to list ten of the most important events in the human record, choose one of the events, and name three other events that would not have occurred or would have been very different if the chosen event had not taken place. In the *Exchange Student* set, students were asked to answer uncomplimentary questions about the United States asked by an otherwise friendly Western European exchange student.

Students were asked to list ten of the most important events in the human record, choose one of the events, and name three other events that would not have occurred or would have been very different if the chosen event had not taken place.

In the *Malbavia* set, students were asked to assume that they were part of a UN team sent into a Third World kingdom undergoing modernization as a result of newly discovered mineral wealth to make observations on the societal effects of these changes and to make recommendations related to norms and traditions such as hospitality and human sacrifice.

Student responses were assessed separately using overlapping criteria embodied in the Assessment Committee's definition of the socioethical understanding. The criteria include:

- A sense of what has been important in history
- A readiness to understand present phenomena in their relation to past causes and future effects

- An understanding that different kinds of factors interact to explain societal phenomena
- An understanding of mainstream American values as the values of a particular society
- Freedom from ethnocentrism
- Freedom from stereotyping
- The ability to provide a reasoned defense of one's own values.

The general results of the assessment showed that upper-division students scored higher than first-year students and that Fredonia students did better than other students who had not had comparable experiences. However, in many cases, both upper-division and first-year students' scores were low on the rating scales, and there were instances of little or no growth. Especially disturbing to the Assessment Committee were students' low capacity and inclination to identify biases and assumptions in their reading and in their own thinking, as well as the lack of improvement demonstrated in problem solving. An issue that needs ongoing consideration is the appropriateness of the expectations reflected in the criteria and scoring scales used to assess the GCP skills.

Increased student writing across all undergraduate majors has resulted from participation in an annual workshop on writing by about 90 faculty.

The assessment results have been used to stimulate a number of initiatives relative to the perceived limited improvement among students. For example, faculty have participated in workshops on developing teaching materials to help students improve thinking and mathematical problem-solving skills. Increased student writing across all undergraduate majors has resulted from participation in an annual workshop on writing by about 90 faculty. There also has been an increased emphasis in Part 3 courses on the study of cultural diversity in America, accompanied by a faculty workshop on this topic.

Of particular interest is a commitment from five departments (Chemistry, English, Foreign Languages, History, and Political Science) to experiment with student portfolios as a way to encourage students to integrate more thoughtfully the goals of the General College Program with the goals of the major.

The FIPSE tests were administered for a second time in the summer of 1991. About 859 incoming first-year students took the battery of tests and will take them again in two years' time. For more information about the Fredonia assessment program, contact: James R. Hurtgen, Dean for Liberal and Continuing Education, W107 Thompson Hall, State University of New York College at Fredonia, Fredonia, NY 14063. Tel: (716) 673-3441.

Assessment Update
March-April 1992
Volume 4, Number 2

Kent State University Regional Campuses

Guest Columnist: William E. Knight.

The Kent State University Regional Campuses constitute a system of seven associate degree-granting institutions located in six northeastern Ohio counties. The system includes 8,944 students and 236 full-time faculty, and shares student and curricular characteristics typical of community colleges. The majority of students are delayed entrants to higher education (the average age is 29); 70% are women; and most are part-time, combining enrollment with work and family responsibilities.

Retention, "stop-out," and intercampus transfer rates are determined through use of the Student Flow Model.

Two types of programs are offered, both baccalaureate-parallel coursework in liberal arts and associate degree technical programs in business, engineering, and health-related technologies, along with a strong developmental education program. Nevertheless, the campuses benefit from their status as part of a comprehensive university with a good academic reputation.

Existing assessment activities were enhanced and systematized, and new initiatives were undertaken for the regional campuses in February 1990 with the establishment of an Office of Academic Assessment and Evaluation Services. The Office has developed a comprehensive assessment plan tailored to the mission and goals of the system.

Assessment Activities

Nine different assessment activities are currently in progress. The goals of access and persistence are assessed through the analysis and distribution of student demographic and enrollment information. Retention, "stop-out," and intercampus transfer rates are determined through use of the Student Flow Model. Comparative data are provided for specific groups such as conditionally admitted students, older students, and women. A wide array of student profiles is provided in the annual *Fact Book*. The *Time to*

Assessment Update
May-June 1992
Volume 4, Number 3

Associate Degree Study was a retrospective analysis of the characteristics and educational experiences of May 1990 associate degree graduates.

Understanding students' educational, career, and personal goals and how they relate to enrollment behaviors are particularly important for two-year colleges. Such goals are routinely determined through systemwide surveys of new freshmen, students who are admitted but do not enroll, nonreturning students, and alumni. Of course, students' goals change over time and maintaining an updated database of students' goals is a challenge for future consideration.

Assessing writing, reading, and mathematics skills for incoming students and documenting developmental course placements are other important goals of the regional campuses. Functions of the Assessment Office include determining placement rates, tracking students who need help in basic skills, and learning about student satisfaction through academic and student services provided at the seven campuses. Evaluations of a large number of services and programs are performed both on systemwide and campus-specific bases. Additionally, results of student evaluations of instruction, which are required in all courses, are disseminated each semester to faculty in order to promote improved teaching and learning.

Assessing writing, reading, and mathematics skills for incoming students and documenting developmental course placements are important goals of the regional campuses.

The University's Liberal Education Requirements Curriculum Committee decided to begin using the ACT College Outcome Measures Program (COMP) Objective Test at the regional campuses on a pilot project basis in 1992. Use of the COMP was one part of a comprehensive effort to evaluate the liberal education program for a North Central accreditation review. A challenge for the assessment program was determining which students to test, how to test those who frequently stop out, transfer, or were not motivated by graduation requirements, and how much value was added by the regional campuses for students who transfer to a senior institution.

The success and persistence of students who continued their studies at a senior institution was assessed through the *Bachelor's Degree Study*. Characteristics and enrollment patterns were analyzed for students who began their studies at the regional campuses and graduated from the main campus with bachelor's degrees in May 1991. Grades, credits to degree, and other factors were contrasted for students who began at the regional campuses, students who began at the main campus, and transfer students.

The success of students who graduate from associate degree technical programs is routinely assessed using job placement and earning studies. Also, student licensing examination results in nursing, radiologic technology, physical therapy assisting, and occupational therapy assisting are examined longitudinally and compared with state norms.

The system's progress toward the goals of economic development and community outreach are monitored through studies of the campuses' contribution to the local workforce, inventories of continuing studies classes and activities offered, and student/client evaluations.

Assessment Update
May-June 1992
Volume 4, Number 3

Following a visit to the campuses by Tom Angelo in 1991, the classroom research model proposed by Pat Cross and Angelo received widespread discussion and was utilized at the campuses.

Use of Assessment Information

The results of assessment activities within the regional campuses receive wide distribution and consideration. Results are shared with the Regional Campuses Faculty Advisory Council, Student Affairs Council, public relations directors, developmental education and central office staff, the president, provost, deans, and other Kent Campus administrators. An Information Officers Advisory Group was formed to help plan assessment activities and consider results. The ERIC Clearinghouse for Junior Colleges sent copies of major studies to over 700 colleges and universities.

Assessment information is used to support planning and decisionmaking within the campuses and at a system level. Enrollment management and other student affairs functions make use of results to monitor and improve their activities. Public relations activities have also been influenced by assessment results, and marketing/advertising surveys have been used at two of the campuses in collaboration with the Assessment Office. The results of assessment efforts were critical for the University's accreditation self-study in 1992-93. Consultants who visited the regional campuses remarked on the assessment program's "informative and well designed reports concerning students and programs."

The results of assessment activities within the regional campuses receive wide distribution and consideration.

Future Initiatives

Following a visit to the Outcomes Assessment Institute at Knoxville, a group of faculty and staff reviewed the campuses' assessment efforts. The use of portfolio review and possible testing in the associate degree major fields were discussion topics. Community image surveys, employer evaluations of graduates, and ethnographic studies were additional assessment projects for future consideration.

The Assessment Office planned to administer the Community College Student Experiences Questionnaire to a broad cross-section of students across the seven campuses in the spring of 1992. Another activity planned by the Office was to receive and analyze information on the success and persistence of students who transferred from the regional campuses to senior colleges and universities. A statewide conference on two-year college assessment hosted by Kent State's Office of Academic Assessment was planned for spring 1993.

Assessment Update
May-June 1992
Volume 4, Number 3

Winthrop College
Rock Hill, South Carolina

The Winthrop program uses multiple methods to assess readiness skills, general education, major studies, and personal growth and development.

Winthrop College, founded in 1886, is a century-old residential public college in South Carolina that emphasizes undergraduate education. Its 5,000 students are enrolled in the College of Arts and Sciences and in the professional schools of Business Administration, Education, and Visual and Performing Arts. Winthrop offers more than 50 liberal arts and professional undergraduate degrees and over 40 graduate degrees.

Winthrop College's assessment effort began in 1986, with a faculty forum on the assessment of student learning. The forum resulted in a plan for a comprehensive assessment program including the establishment of an Office of Assessment to coordinate program development and implementation. The assessment program was shaped by input from faculty and student advisory boards. These boards continue to influence the direction of the assessment program, which is intended to provide information that can be used to improve college programs and services.

Students are encouraged to take an active part in the assessment process. A student guide to the assessment program asks students to consider such questions as these: Are you acquiring the knowledge necessary to be an educated adult in today's society? Are you gaining the skills you need to have a successful career within your chosen field? To what extent are you developing personally as a result of the diverse experiences offered at Winthrop? Are you getting what you need and want in an education?

The Winthrop program uses multiple methods to assess entry-level and college readiness skills, general education, major studies, and personal growth and development. These methods are designed to go beyond the grading conducted in individual courses to obtain measures of learning and development throughout the course of a college education. For example, the results of tests of basic knowledge and skills and of placement tests in writing and mathematics are used in combination with high school performance indicators and college entrance examinations to determine students' readiness for college.

Assessment Update
July-August 1992
Volume 4, Number 4

The general education program at Winthrop College focuses on such basic skills as written and oral communication, quantitative methods, critical thinking, and international understanding. These diverse skills are assessed through various standardized tests and through written essays and classroom assessment techniques being developed by faculty.

General education assessment has four parts. The first is a timed essay and the institutional matrix of the College BASE (a criterion-referenced test of general education marketed by Riverside Publishing Company), which are administered to freshmen and seniors in the fall of alternate academic years. These tests provide information on cognitive abilities and specific subject matter knowledge and skills, which is fed back to departments to help them assess their programs.

The second part of the general education assessment, the Student/Alumni Follow-Along and Career Tracking System (FACTS), gathers information on respondents' perceptions of the quality of the general education curriculum and of instruction for ongoing career planning and development. A Senior Survey (a computer-assisted telephone interview conducted 6 to 12 months after graduation), an Admitted Student Questionnaire (developed and scored by the College Board), and a Withdrawing Student Survey provide information about the quality of the general education curriculum, evaluate career preparation efforts, and enhance the attainment of realistic career goals for Winthrop graduates.

> *Assessment in major studies areas requires that each academic unit at Winthrop assess the attainment of educational goals and objectives.*

Using input from faculty on the types and amounts of writing that students actually do, an interdisciplinary Writing Assessment Committee developed the third part of the general education assessment, a Class Portfolio model designed to evaluate curricula rather than students. The committee has also developed writing evaluation criteria that blend holistic and analytic scoring and has created ways to stimulate writing.

The fourth general education assessment, the Learning Research Project, was designed to "positively affect the fulfillment of general education goals by providing individual instructors with immediate and relevant information on their students' learning." The program is intended to encourage and facilitate faculty-controlled classroom "assessment-for-improvement" efforts.

Assessment in major studies areas requires that each academic unit at Winthrop assess the attainment of educational goals and objectives. The methods selected or developed are appropriate for the goals and objectives set for each major. For example, laboratory skills assessment is used in biology, portfolio review in art, written and oral communication skills assessment in modern languages, and classroom teaching skills assessment in education.

The final assessment effort concerns student growth and development. Begun in 1989-90, this effort has to date resulted in definition of the characteristics of student growth and development: social responsibility (encompassing altruism and integrity), appreciation of diversity, maturity, and self-esteem. In addition, during 1991-92, 600 freshmen were given the Learning and Study Strategies Inventory (LASSI), 150 freshmen were given the Cooperative Institutional Research Program (CIRP) survey, and a random sample of freshmen and seniors were given a survey to determine their level of involvement in social issues. The Division of Student Life and Library Services also documented the resources available to students and analyzing how these resources facilitate students' personal growth and development. Another result of this effort has been the development of a *Student Growth and Development Annotated Bibliography*.

Ten general principles guide the assessment of student learning and development at Winthrop College:

- The primary reason for assessment is to improve student learning and development.
- The assessment program is designed primarily for internal use in making decisions that seek to improve programs, instruction, and related services.
- Assessment program initiatives must include training and related resources for faculty and student support personnel who are responsible for assessment activities.
- Participation by faculty in student assessment activities will be appropriately supported and recognized by the college.
- The development of an effective, valid assessment program is a long-term, dynamic process.
- The assessment program will seek to use the most reliable, valid methods and instruments of assessment.
- Assessment must involve a multi-method approach.
- The technical limitations of the respective data will be considered during subsequent decisionmaking and delineated in assessment reports.
- Assessment results are not intended to be used punitively against students.
- Assessment of student learning and development is a process distinct from faculty evaluation.

Within the context of these principles, Winthrop College appears to be developing a comprehensive assessment program focusing on readiness skills, general education, major studies, and students' personal growth and development. For more information about the current status and impact of this program, contact Joseph Prus, Associate Professor of Psychology and Director of the Office of Assessment, Winthrop College, Rock Hill, SC 29733. Tel: (803) 323-2341.

Ten general principles guide the assessment of student learning and development at Winthrop College.

Assessment Update
July-August 1992
Volume 4, Number 4

Capital University

Guest Columnists: Richard A. Schalinske, associate director, Assessment Center; Robert A. Patterson, director, Assessment Center; and Gary L. Smith, dean, Adult Learning and Assessment, Capital University.

Capital University is a four-year, liberal arts-based Lutheran institution located in Columbus, Ohio. It is organized into four undergraduate units, including the College of Arts and Sciences, the Conservatory of Music, the School of Nursing, and the Adult Degree Program. In addition, there are two graduate colleges: the Law and Graduate Center, and the Graduate School of Administration. The university offers six undergraduate degrees, five graduate degrees, and more than thirty undergraduate majors to its approximately 3,400 students.

The Assessment Program

Five years ago, Capital University began developing a program to assess a new competency-based core curriculum and the general university mission. It was assumed that a well conceived assessment program would empirically verify many of the educational assumptions made by the institution. For example, the mission states that students will experience academic, personal, and social growth through the core courses and over their four years at the university. Therefore, the assessment program should measure changes not just in academic proficiencies but also in student attitudes and values.

> T*he assessment program should measure changes not just in academic proficiencies but also in student attitudes and values.*

From the outset, the University Assessment Center assembled an integrated, multiple-measure Assessment Battery that included the Shipley Institute of Living Scale, a measure of intellectual ability for use in group comparison studies; the Canfield Learning Style Inventory, a measure of preferred learning conditions and styles for developing individual and group profiles; the Jackson Personality Inventory, a measure of personality traits and styles for developing individual and group profiles of attitudes and values; and the Academic Profile, a measure of academic proficiency for developing

Assessment Update
November-December 1992
Volume 4, Number 6

BEST COPY AVAILABLE

individual/group profiles and comparisons with normative scores and criterion-referenced proficiency levels.

Current Assessment Procedures

All incoming freshmen are administered the Assessment Battery during summer orientation. A computer-generated report that describes students' learning styles in terms of their significant Canfield scales and the student report form from the Academic Profile are distributed to the appropriate academic advisers. Advisers discuss the reports with their advisees and use the information to help plan academic programs. Students requesting a more detailed interpretation of their results are referred to the Assessment Center, where they receive an expanded custom report that contains all of their scale scores from the Canfield, Jackson, and Shipley instruments.

Once individual student assessment data have been reported, the overall results are analyzed and integrated in a variety of ways. Initially, a serial "white paper" is distributed to the campus community. Each edition uses one page of graphics and jargon-free text to describe the current class of freshmen. Later a more detailed analysis is conducted and distributed, with comparison by major, college, age, and gender.

Data from each class are stored for use in the longitudinal assessment of the university mission.

Findings from these studies are also shared with local educators. In a recent study on gender differences in learning styles for current incoming freshmen, published results were found to be applicable not only at the university but also at a neighboring private school for girls.

The Student Survey, a qualitative instrument with structured and unstructured question sets, was developed for constructing subjective group profiles, corroborating qualitative data, and conducting longitudinal group comparisons. The Student Survey is administered to freshman students after they have been at Capital University for three months. An equivalent form, with minor wording changes, is administered to seniors at the same time. The results are used to provide basic information about the incoming freshman class and to compare results from the same class of students on the senior form four years later. Although the groups are not equivalent, a tentative, interim report comparing freshman and senior perceptions suggests changes that may have taken place after four years at the university.

Uses of Assessment Information

Data from each class are stored for use in the longitudinal assessment of the university mission. In this annual study, an institutional assessment question is posed: was the university successful in accomplishing its stated mission for one class of students? More specifically, did Capital University provide "quality higher education," as evidenced by enhanced academic proficiencies or student perceptions? Did Capital help students "prepare

Assessment Update
November-December 1992
Volume 4, Number 6

themselves for personally rewarding lives" in terms of careers and relationships? Did Capital graduates exhibit a "strong desire for lifelong learning"? Were the faculty and staff perceived as being "outstanding"? Did graduates exhibit more "mature" attitudes and values? When Assessment Battery results from the freshman and senior years for the same class of students are compared, these questions can be answered in a manner that is empirically sound and suitable for administrative consideration.

In the first assessment study for the class of 1988, the claim that Capital University provided "quality" higher education was supported both by evidence of enhanced academic proficiencies in reading, writing, mathematics, and critical thinking and by the favorable self-report of perceptions on the Student Survey. The claims that Capital would prepare students for "personally rewarding lives" and that graduates would exhibit more "mature" attitudes and values were also supported by qualitative and quantitative findings suggesting positive personality and learning-style changes had occurred over four years. These findings also supported the contentions that Capital graduates would exhibit a desire for lifelong learning and that the faculty and staff would be perceived as "outstanding." Thus, the university had probably fulfilled much of its mission for the class involved in the study.

The claim that Capital University provided "quality" higher education was supported both by evidence of enhanced academic proficiencies in reading, writing, mathematics, and critical thinking and by the favorable self-report of perceptions on the Student Survey.

Future Directions

Having established the basis for an ongoing assessment of the university mission, the Assessment Center is now refining the existing program and expanding it into more specific areas. The assessment program continues to flow from a macro or global perspective to more micro or practical perspectives. A mandatory student testing policy has been added, and family surveys and focus group interviews for freshmen and seniors are being considered as additions. If these refinements are added, Capital University should be able to assess the achievement of its mission with greater precision and clarity.

The Assessment Center is currently working with faculty to develop practical, valid procedures for assessing proficiencies in each of the core curriculum courses and, eventually, in each academic major. By integrating objective and subjective data from these macro and micro perspectives, the center will further enhance an already credible assessment program.

For more information about the assessment program at Capital University, contact one of the authors at Capital University, Room 030, Renner Hall, 2199 E. Main Street, Columbus, OH 43209. Tel: (614) 236-6284.

Midlands Technical College

"Cutting edge" legislation mandated that effective systems of quality assessment and accountability be established and maintained by South Carolina public colleges.

Midlands Technical College (MTC), which serves the Columbia, South Carolina, metropolitan area with three campuses, has seen a 61% enrollment growth since 1978. It currently has over 11,000 students, enrolling nearly one-quarter of the area high school 1990 graduating seniors who went on to college. MTC also serves approximately 18,000 area residents through its Continuing Education Division.

In 1986, the faculty, staff, and governing board of MTC initiated a strategic planning process, which resulted in the document Vision of Excellence. This document provided the foundation for a comprehensive institutional effectiveness initiative begun in 1988.

At this same time, South Carolina enacted a law known as the "cutting edge" legislation. This law mandated that effective systems of quality assessment and accountability be established and maintained by South Carolina public colleges. Systems must be designed to determine institutional effectiveness, disseminate the results of outcomes to constituents within the state, and initiate changes in curriculum, programs, and policy based on data related to institutional effectiveness. The MTC initiative goes beyond South Carolina's mandated requirements, and, in 1989-90, MTC received the designation of lead institutional effectiveness college among South Carolina's two-year technical and community colleges.

In defining institutional effectiveness, MTC also goes beyond the traditional definitions of program evaluation and assessment of student outcomes. MTC considers it to be an institutional perspective that focuses on accurate planning, assessment of accomplishments (both of students and the institution's overall effectiveness), and the use of assessment results to plan and make decisions. Institutional effectiveness has two major components: planning, which is defined as a process of documenting the intended purpose, direction, and expected outcomes of the college and providing foresight in the formulation of policies, programs, and services; and evaluation, which is

Assessment Update
March-April 1993
Volume 5, Number 2

defined as a process of measuring the college against its stated purpose and indicators of effectiveness in terms of outcomes accomplished.

The planning process involves both strategic planning and operational planning. Strategic planning is used to identify the major direction and priority initiatives of the college. This begins with clarification of the college's mission, its role and scope, and its values. In order to achieve realistic results, MTC is clear about the internal and external contexts within which the college operates.

Based on these considerations, a set of eight institutional goals was created for the period 1992-97. These goals include enhancing and developing the curricula to meet multiple challenges; providing the highest quality instruction through excellence in teaching and comprehensive instructional support; and maintaining and refining support processes that enhance student success. In turn, each goal is associated with a set of priority initiatives, such as the implementation of new general education core requirements and implementation of the instructional component of effectiveness measures, including program review, needs analyses, and productivity measures.

While long-range goals and related initiatives are necessary to give an institution direction, they are not sufficient to provide a complete picture of ongoing programs and services and their effectiveness. For this purpose, MTC developed a set of assessment criteria to answer the question, "How effective is our institution in providing ongoing programs and services that encourage student success and support our mission?" These criteria are called "critical success factors" and include accessible and comprehensive programs of high quality, student satisfaction and retention, post-education satisfaction and success, economic development and community involvement, sound and effective resource management, and dynamic organizational involvement and development.

> *While long-range goals and related initiatives are necessary to give an institution direction, they are not sufficient to provide a complete picture of ongoing programs and services and their effectiveness.*

According to MTC documents, critical success factors are the key things that must succeed for the organization to flourish and achieve its goals. They are defined in a way that guides the development of indicators of effectiveness and sets of measurable criteria in response to two questions: "What do we want the results of our college's effectiveness to be?" "What specific evidence are we willing to accept as an indication that the results have actually been achieved? For example, the following are two critical success factors and their related indicators of effectiveness: (1) student satisfaction and retention: accurate entry testing and course placement, retention to achievement of students' goals, satisfaction with instruction and personal growth, and assessment of student services; and (2) post-education satisfaction and success: graduate employment and continuing education, employer satisfaction with graduates, and alumni satisfaction with education and training.

The indicators and their supporting measurement criteria are the observed, quantified, and qualified results of performance on the critical success factors. For example, two of the standards set in relation to the student satisfaction and retention factor and one of its indicators, retention to achievement of students' goals, are as follows: retention of subpopulation groups will be within 5% of the college-wide average, and the freshman-to-sophomore retention rate will be at or above the national retention rate for two-year public colleges. The factors, indicators, and standards (together with their results) are reviewed each year to provide an update on the overall institutional effectiveness of MTC.

Operational planning involves the year-to-year process of converting the strategic plan into action. The operational plan includes the collection of data related to each critical success factor. Sources of information include the student database on enrollment, retention, race, gender, grade point average, and developmental performance; program enrollment, full-time equivalents, section sizes, and grade point averages, graduation rates, competence performance, placement rates, senior projects, licensure test results, transfer performance, advisory committee reviews, and instructor loads; surveys of current students, withdrawing students, graduates, employers, faculty, staff, and community members; support services involving student extracurricular activities, career and counseling services, usage reports, and financial aid reports; administrative data on cost per full-time equivalent, financial audits, college goals, facilities usage studies, and external funding; and comparison to normed data from standardized tests, licensure examinations, and national standards and guidelines.

Those making decisions at all levels would have the best information possible and they would use the information in support of student success and the college's mission.

The evaluation process is an ongoing part of implementing the operational plan. Of course, the preparation of budgets is influenced by the planning process, as well as by anticipated continuing costs.

As the result of a reorganization in 1989 and a merger in 1990, the Office of Institutional Effectiveness and the Office of Research and Analysis were combined, and their new mission was described. This merger helped to ensure that those making decisions at all levels would have the best information possible and that they would use the information in support of student success and the college's mission.

From their experience, MTC faculty and administrators have identified the following principles for implementing institutional effectiveness programs:

- Secure support from top leaders and keep them informed and involved. Connect assessment to the college purpose and make it part of a campus-wide systematic planning process.
- Create internal motivation by deciding which questions about student learning and the college's programs and services are most important to answer, since not everything can be assessed.
- Assess where the institution is and what is currently being done first, and then build on that.
- Review and study what others have done so as to avoid reinventing the wheel.

Assessment Update
March-April 1993
Volume 5, Number 2

- Develop or obtain technical expertise as needed and clearly assign responsibilities.
- Organize the effectiveness assessment so that it encourages active involvement at all levels in collaborative efforts and in data gathering processes.
- Determine how effectiveness will be assessed using creditable research methods and multiple measures.
- Focus on outcomes, performance, effective communication of results, and their use for decisionmaking. Follow up to ensure that results produce change.

For more information about the institutional effectiveness effort at MTC, contact Dorcas A. Kitchings, Director, Research and Analysis, Midlands Technical College, P.O. Box 2408, Columbia, SC 29202. Tel: (803) 738-1400.

Mt. Hood Community College

Assessment of its effectiveness has been part of Mt. Hood Community College's operation from its earliest years.

Assessment Update
May-June 1993
Volume 5, Number 3

Mt. Hood Community College, which serves the suburban population east of Portland, Oregon, typically has 30,000 individual students enrolled over a year's time in courses that are designed for exploration and self-improvement as well as those that lead to certificates and degrees. Noncredit enrollment makes up 43% of the total course enrollment and was up nearly 20% in the 1991-92 academic year. Over one-third of Mt. Hood's students enroll in evening classes, and nearly 30% are over 30. Hispanics are the fastest growing segment of the population.

Assessment of its effectiveness has been part of Mt. Hood Community College's operation from its earliest years, but formal assessment activities have increased over the last ten years. Beginning about five years ago, faculty and staff became involved in such efforts as strategic planning, program improvement, and budget development. Student success and a related focus on teaching and learning have been topics of special emphasis in these activities.

The purpose of these changes has been to establish processes and practices that support the establishment of "an integrated systems approach to organizational effectiveness." This concept, first described by Argyris (1970), established the importance of valid information made freely available to the organization (for example, the system), since without such information those who are part of the organization cannot make informed choices, commit to a course of thinking or action, or feel any ownership of processes or outcomes. Argyris's ideas led to the notion of "the learning organization" (Senge, 1990)—that is, organizations in which people, through knowledge, understanding, and learning, increasingly expand their capability to shape the future. Such an organization is built on the principle of an integrated system in which all feel as though they understand and are a part of the whole organization. In contrast, if people understand only a part of the organization and, as a result, practice departmental separatism, follow management trends for their own sake, and preserve top-down thinking, the result is likely to be transient successes and demoralized, uninvolved staff.

BEST COPY AVAILABLE

Mt. Hood Community College's institutional effectiveness activities are guided by the ideas of Argyris and Senge and by the following questions: (1) What is our purpose? (mission/mission themes); (2) What do we do to fulfill our purpose? (practices/processes); (3) What results do we expect when we do what we do? (expected outcomes); (4) How do we know when we have achieved results? (performance); and (5) How do we know when our results are effective? (measurement).

With the establishment in the 1980s of student success as the primary institutional purpose, Mt. Hood began a process of continuous instructional program improvement based on strategic planning and assessment tied directly to budget development. There are four primary parts to what Mt. Hood calls an integrated system of effectiveness.

The first part is the assessment of student success in relation to student intentions (the overall mission of the institution). Because of the diversity of incoming students, it is not possible to assume a narrow definition of student success or student intentions. Therefore, a mandatory assessment and placement program called Guided Studies was created in 1984. Through this program, data are collected from students when they begin their studies regarding their basic skills and their reasons for attending college. Then their progress is monitored over time, both while they are in school and after they have left the college for employment related to their course of study or for further study at four-year institutions. Individual departments also have devised a variety of strategies for enhancing and assessing student progress that range from course-specific placement tests to competency-based outcomes measures.

With the establishment of student success as the primary institutional purpose, Mt. Hood began a process of continuous instructional program improvement based on strategic planning and assessment tied directly to budget development.

Program improvement is the second part of Mt. Hood's integrated system of effectiveness. This process is based on 19 student outcomes factors: student demand, job placement and transfer success, employment outlook, instructional cost effectiveness, facility requirements, revenue projections, course retention, student success by course, retention in sequential courses, discipline/program retention, completion, staff development, quality of the curriculum, service to students from other disciplines, instructional alternatives within the college, instructional alternatives at other institutions, comprehensive/balance, instructional offerings, and service to the community. These qualitative and quantitative factors are reviewed annually; if problems in program effectiveness are identified, an in-depth review is conducted.

The information gained in the program improvement process is used in the third part of Mt. Hood's integrated system of effectiveness—that is, the strategic planning and resource allocation (budget development) process. Each year a series of collaborative meetings are held that involve faculty, support staff, management, and student leaders. External factors are incorporated into this process through input from program advisory committees, institutional partners in the community, and environmental activities. This information is incorporated into each year's action plan, which includes a reiteration of the institutional philosophy, an analysis of current conditions, and a description of changes that will be implemented given available resources.

The interrelated activities of strategic planning, budget development, strategic plan assessments, and program improvement form a dynamic and ever evolving approach to renewing institutional vision and solving institutional problems. A key element of this approach is the involvement of the Strategic Planning Council, which represents all groups on campus. The council proposes institutional priorities, which are fed into the budget development step. In addition, the council coordinates the strategic plan assessments (including collecting data on the program effectiveness factors) that occur each spring. In this step, each area of the college evaluates its progress in meeting its program improvement objectives.

Based on the results of the strategic plan assessments, the various academic and support areas create new program improvement plans. In many cases these involve staff and organizational development, the fourth part of Mt. Hood's integrated system of effectiveness. Student success depends on faculty and staff who are enthusiastically involved in the continuous learning and development of each student. Student success is positively influenced by the quality and intensity of student-faculty interaction. To facilitate student success, faculty need staff support, information, and appropriate technology. Both they and other staff members must feel self-confident, purposeful, and valued by the institution.

To facilitate student success, faculty need staff support, information, and appropriate technology.

Staff and organization development, based on continuous self-assessment, is an integral part of the system of effectiveness and includes a structured program of teaching improvement, a Small Group Diagnosis program, and classroom research based on the ideas of Cross and Angelo (1988). Faculty are involved in an intensive two-week program of learning activities prior to the beginning of classes each fall. Plans also are being developed for a teaching and learning center (TLC) that will provide learning opportunities for faculty, administrators, and support staff.

By creating an integrated effectiveness system connecting the institutional mission with the processes of strategic planning, program improvement, and faculty/staff development, Mt. Hood has created a model for building faculty and staff commitment to continuous improvement, which ultimately leads to accomplishment of the institutional mission and to improved student success.

For more information about the assessment program at Mt. Hood contact: R. Dan Walleri, Director of Research, Panning, and Computer Services, Mt. Hood Community College, 26000 S.E. Stark Street, Gresham, OR 97030, (503) 667-7146.

References

Argyris, C. *Intervention Theory and Method: A Behavioral Science View.* Reading, Mass.: Addison-Wesley, 1970.

Assessment Update
May-June 1993
Volume 5, Number 3

Cross, K. P., and Angelo, T. A. *Classroom Assessment Techniques*. Ann Arbor: National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan, 1988.

Senge, P. "The Leader's New Work: Building Learning Organizations." *Sloan Management Review*, 1990, 32(1), 7-23.

Assessment Update
May-June 1993
Volume 5, Number 3

Union College

"The academic program is designed to meet the College's mission of providing its students with the opportunity to increase their general knowledge and understanding of the world around them and to obtain the skills necessary to compete in the job market."

Union College is an independent liberal arts institution located in Barbourville, Kentucky. It has been affiliated with the United Methodist Church for over 100 years. Central Appalachia is Union's primary service area and home to the majority of the college's nearly 1,000 students. Through technical and skills-oriented programs, Union helps to improve the human resources of the region and, in particular, to meet the personnel needs of the public school systems and the ministry. Besides serving as a cultural center for the community, Union also offers Appalachian Studies programs for local students and those visiting from other institutions with the purpose of educating people about Appalachian life, traditions, and problems.

Stimulated by the new initiative of the Southern Association of Colleges and Schools, Union College decided to develop an assessment program. This decision led to the obvious conclusion that such a program must flow from a clear mission statement, as well as specific goals. The Planning and Evaluation Committee, appointed by the vice president for academic affairs, proposed a revised mission statement and a set of related goals for general education during the 1987-88 academic year. The committee consisted of the director of planning and evaluation, faculty from all academic divisions (music, history, business management, environmental science, and sociology), and the director of special services (representing administration and staff).

The college's academic mission is summed up in the following statement of purpose: "The academic program is designed to meet the College's mission of providing its students with the opportunity to increase their general knowledge and understanding of the world around them and to obtain the skills necessary to compete in the job market." In particular, "the basic undergraduate liberal arts curriculum makes available to students a large body of knowledge drawn from the humanities, natural sciences, and social sciences. [Studies] in pre-professional, technical, and skills-oriented fields enhance postgraduate employment opportunities" (Union College Board of Trustees, 1988).

Assessment Update
September-October 1993
Volume 5, Number 5

The statement of purpose, or mission statement, provided a context for three types of goals. One type focused on the basic undergraduate liberal arts curriculum, which affects all undergraduate students. The second was set by each department and administrative office, as well as by special service and support programs. The third type of goal was created in relation to the following dimensions of the mission statement: communication skills, ethics and values, general knowledge, critical thinking and problem solving, context of modern civilization, and learning skills and personal development.

Revision of the College mission statement and the development of a set of related goals for general education constituted the first phase of the Union College Assessment Program.

The second phase of the program involved the planning and implementation of an assessment plan for the eighteen general education goals. For example, one of the major goals of Union's liberal arts curriculum is to help students develop and improve their communication skills. This goal is refined into four behavioral statements that indicate what a student should be able to do regarding both expression and comprehension related to communication, such as: write a clear and cogent essay and prepare and deliver a logical and organized oral presentation.

To ensure faculty input, a faculty survey was used to assess each goal.

The ethics and values goal is associated with statements to the effect that the Union student should be tolerant toward and respectful of these different views and ideas and able to think about and articulate problems and issues related to ethics and values.

Each goal was assessed using multiple measures. The measures included several locally developed assessment instruments, commercial tests, and existing records. To ensure faculty input, a faculty survey was used to assess each goal. Other locally developed instruments included a survey of student use of facilities, such as the library, and an alumni survey focused on personal development.

Locally developed senior interviews and senior essays were used to assess such objectives as ability to think about and articulate problems and issues related to ethics and values (interview) and ability to think and reason in a critical, logical, and objective fashion (essay).

The subscores obtained from the standardized Academic Profile test were used to assess such areas as writing, humanities, natural and social sciences, mathematics, critical thinking, writing, and reading. Another commercial instrument used was a survey of students' self-estimation of gain. This survey provided information on accomplishment of the goals related to communication skills, general knowledge, and the context of modern civilization.

Extensive use was made of existing information gathered as part of the regular grading process—primarily grades in core courses such as English, sciences, history, and religion. In addition, student performance in core lecture courses was used to assess students' ability to listen to a lecture or presentation and recall and understand the major points. Student performance and attendance also were used to assess ongoing interest in and appreciation for music, literature, and the arts and the traditions on which these fields are founded. A comparison of student performance in developmental courses and in later courses was used to assess learning skill development.

At Union College, many specific recommendations resulted from the first two assessment cycles of the general education curriculum. One of the most important was the decision to conduct these assessments every other year. This decision was made because other assessments must be designed and conducted (for example, of majors, minors, and programs) and because time is needed to develop thoughtful responses to the assessment information already collected.

In addition, changes were made in facilities, with the addition of a computer laboratory; in curriculum, with the addition of world religions and political philosophy courses; in staff, with the addition of two foreign language teachers; in recruitment of students from outside the immediate region; and in several other areas.

The Union College Assessment Program is an important means of identifying the institution's strengths and weaknesses so that the former can be enhanced and the latter can be corrected.

Clearly, the Union College Assessment Program is not just a response to new criteria established by the Southern Association of Colleges and Schools. It is also an important means of identifying the institution's strengths and weaknesses so that the former can be enhanced and the latter can be corrected.

For more information about the Union College Assessment Program, contact Paul S. Moore, Vice President of Academic Affairs, Union College, 310 College Street, Barbourville, KY 40906. Tel: (606) 546-4151.

Reference

Union College Board of Trustees. "Statement of Purpose (Mission Statement)." In *Union College Status to Report on the College's Plan to Comply with Section III: Institutional Effectiveness.* Barbourville, Ky.: Union College, 1988.

Assessment Update
September-October 1993
Volume 5, Number 5

PETER J. GRAY

Eckerd College

In 1986, Eckerd College, a small liberal arts institution in St. Petersburg, Florida, established a special faculty task force charged with developing a comprehensive set of policies to guide the systematic evaluation of the college's academic programs and other closely related activities. The task force conducted extensive hearings on the creation of an assessment program. Participation in the hearings gave faculty, administrators, and students an opportunity to understand and identify with the need for assessment and to develop a sense of ownership in the plan that emerged.

The assessment of general education has received the most attention.

Major elements of the plan were the creation of an educational assessment coordinator position and a standing faculty committee. Since 1987, the committee and the coordinator have worked together with the dean of faculty and the general faculty to implement a comprehensive assessment program at Eckerd.

The program has several key components, including the assessment of majors, student residential life, general education, and international education. However, the assessment of general education has received the most attention.

Assessing the college's general education program presents special challenges because many of the courses in this program are interdisciplinary and unique to the college. At present, there are no commercially available tests that the assessment coordinator and faculty committee feel appropriately measure the development in knowledge and values expected in the Eckerd College general education program.

In response to this challenge, the approach at Eckerd College has been to develop a set of questionnaires for freshmen, seniors, alumni, experienced learners, and faculty. For example, the Entering Student Survey asks students about their expectations regarding the influence of Eckerd College on such areas as skill with computers, life values, writing ability, religious beliefs, personal integrity, aesthetic sense, ability to locate information, enjoyment of learning, confidence in their future, speaking ability, and

Assessment Update
November-December 1993
Volume 5, Number 6

openness to others. Students are asked to indicate the expected influence of an education at Eckerd College in relation to 40 such items using a scale of 1 to 5 (where 1 equals very negatively, 3 equals neither negatively nor positively, and 5 equals very positively). These items cover a broad range of potential outcomes.

The same items appear on the Senior Survey in relation to the statement "Please indicate the influence that you think Eckerd College has had on you." In addition, the Senior Survey asks students to indicate the extent to which they agree or disagree with such statements as "My choice of Eckerd College was a good one," "My professors were helpful to me," "The program in my major was a good one," and "The campus housing was good." The Senior Survey also contains sections on new graduate information, extracurricular activities, physical activities, and other information for the Alumni Office.

Four surveys provide a database of consistent and comparable information on students' perceptions of the influence of an education at Eckerd College over time.

Recent graduates are sent a survey with the same expanded set of items that make up the Senior Survey, as well as an item about the Alumni Office. The Recent Graduates Survey asks for information about graduates' current activities (for example, graduate study and employment status) and asks them to update their addresses and telephone numbers as well.

Non-traditional age students at Eckerd are part of the Program for Experienced Learners. They receive an extensive survey when they graduate. This survey concerns their personal and family situation, educational and work plans, life changes during the program, and open-ended items about the program and its impact on them. The items on the Entering Student and Senior surveys regarding the influence of Eckerd College and student satisfaction are also part of the Experienced Learners' Survey, with some additions and modifications to make them relevant to this special population.

These four surveys provide a database of consistent and, therefore, comparable information on students' perceptions of the influence of an education at Eckerd College over time. They also ask students to provide feedback on how to improve the general education program and other aspects of the college.

Another group of surveys has items tied directly to the stated course objectives of the general education program courses. There are surveys for the two Western Heritage courses, the lower-division Perspective courses (aesthetic perspective, cross-cultural perspective, environmental perspective, and social relations perspective), the senior-level Judeo-Christian Perspectives course, and the Senior Seminar. Similar surveys are designed for students and for the faculty who teach these courses.

By surveying faculty, in addition to present and past students, a variety of perspectives on the outcomes, current quality, and future direction of the general education program is attained. The college has also found the locally developed surveys to be efficient, quantifiable, stable over time, and inexpensive to administer, process, and summarize.

Assessment Update
November-December 1993
Volume 5, Number 6

In addition to these opinion surveys, which focus on the general education program, there are experiments with discipline-specific comprehensive examinations such as the Graduate Record Examination subject tests, with the Educational Testing Service Academic Profile instrument (administered to freshmen and rising juniors), with writing portfolios for assessing writing competence, and with focus group and entrance and exit interviews of students. By developing these various methods, the college hopes to avoid basing judgments of success or failure only on the opinion of students and faculty, and the tendency to reduce the often subtle and long-term outcomes of general education to short-term, measurable outcomes.

The findings from the various assessment activities are summarized and distributed to the campus community by the coordinator of educational assessment, most often in the publication The Eckerd Profile. In addition, the findings are submitted to the Educational Policy and Program Committee of the faculty, which is responsible for formulating academic policy.

The assessments have shown that students have sharpened their analytical skills, developed a greater sensitivity to value issues, and developed a strong awareness of their cultural heritage. Revisions to the general education program based on findings from the assessments are currently being considered. These include filling gaps shown to exist, including the need to develop students' scientific and aesthetic literacy as well as the need to increase their awareness of the non-Western world and the issues confronting a truly global society.

At Eckerd College, faculty fears of assessment have been reduced significantly.

The findings also have led to several realizations: the various parts of the undergraduate program must be linked more coherently, regular review of discipline programs should be scheduled, and there needs to be a focus on how students learn, not just on learning outcomes.

Since all faculty teach the Western Heritage and other general education courses on a regular basis, they have direct input into the assessment and a stake in the continual improvement of these courses. Faculty are encouraged to experiment with the many practical ideas for improvement that are provided by the survey information, such as active learning, group learning, special projects, and student writing.

By involving students and faculty directly in the assessment process, by taking a gradual approach to building the assessment program, and by letting the faculty take the lead in determining how its results may best be used to improve teaching and learning at Eckerd College, faculty fears of assessment have been reduced significantly. However, it is well to remember that this success has come at considerable cost to the college, including the expenditure of funds related to the assessment position and the time and effort that faculty spend on assessment and subsequent instructional improvement.

For more information about assessment at Eckerd College, contact Lloyd W. Chapin, vice president and dean of faculty, Eckerd College, P.O. Box 12560, St. Petersburg, FL 33733. Tel: (813) 867-1166.

Assessment Update
November-December 1993
Volume 5, Number 6

Virginia Military Institute

There have been several cases in which assessment information has really "made a difference in the way the VMI does business."

The assessment program at the Virginia Military Institute (founded in 1893 as the nation's first state military college) is based on the premise that in order for an assessment program to thrive and make a difference, it must be part of the planning-evaluation-resource allocation cycle. This lesson was learned the hard way. The Virginia Military Institute (VMI) program was stimulated originally by a 1985 state-mandated assessment initiative that focused on assessing student outcomes, and therefore put the cart of assessment before the horse of institutional purpose. That initial mistake was eventually fixed by taking a fresh look at the assessment information collected over the first three years of the program and determining how to make it useful to those involved in assessment.

The rethinking of the VMI assessment program fostered the rediscovery of four basic assumptions, inherent in every VMI program, by which programs are implicitly evaluated. These assumptions are (1) a sense of justice, (2) an adversative model of education, (3) a constructed environment, and (4) candor and openness. The ideal consequence of this set of assumptions is that "a strong leader emerges from this educational mix (RiCharde, Olney, and Erwin, 1993, p. 181). Therefore, leadership is a key educational outcome at VMI.

Assessment personnel identified several developmental assumptions that reflect the basic values of VMI in order to take the next step toward creating a context for the collection and use of assessment information. These assumptions include, first, that VMI student development occurs in periods of rapid change punctuated by plateaus (discontinuous development); second, that students appear to develop in the same general direction (unitary development); and third, that new learning builds on prior learning (cumulative development). The cognitive and affective measures that make up the assessment program are intended to test these assumptions.

The last step in the assessment program development process was to revise the institutional long-range plan around key educational areas called

Assessment Update
March-April 1994
Volume 6, Number 2

competencies-factors that were felt to contribute to the development of leadership as defined at VMI. Areas identified were ability to communicate effectively, ability to think critically, commitment to lifelong physical health and strength, mastery of a major field of study, ability to succeed in an organizational setting, skill in interpersonal relations, and commitment to ethical inquiry and standards of integrity (*Third Assessment Report*, p. 14). In the long-range plan, each of these seven areas was arranged in a hierarchy of educational complexity beginning with the general area and proceeding to cognitive and affective domain manifestations, levels of skill, educational objectives, and, finally, measurable learning objectives. Educational strategies also were attached to each area.

The initial three-year period, which led to the collection of important information about student outcomes (student development, retention, and achievement), may have been needed to stimulate the changes that eventually resulted in a well thought out institutional assessment plan. However, if the program had to be implemented anew, it is probable that efforts to build a program from the 'grass roots' would prove a less painful method of imposing accountability on a college faculty.

As currently configured, the model for the planning and review process at VMI begins with the college mission, which is directly translated into a comprehensive strategic plan. The strategic plan influences several other plans, including the long-range academic plan, the co-curricular plan, the athletic plan, the master facilities plan, and the strategic marketing plan. The long-range academic plan influences departmental or program goals, which in turn influence curricular or program structure and faculty and staff activities. Together these last two elements influence student performance outcomes, that is, student development, retention, and achievement, which are the focus of the assessment program. This assessment information provides the basis for program evaluation and, eventually, for modification of the comprehensive strategic plan.

During the recent history of the program—that is, after the redesign of the VMI assessment program based on the experience of the first three years—there have been several cases in which assessment information "had an impact" (RiCharde, Olney, and Erwin, 1993, p. 180). A number of problems have been identified by the assessment information and corrective measures have been taken.

A number of problems have been identified by the assessment information and corrective measures have been taken.

One of the most important and complex of those problems was that of freshman retention. The problem was manifested in three ways. First, it resulted from the stress of making the transition from high school to a college such as VMI, with its strong military tradition. Second, the stress related to the change in learning style all students needed to make to adjust to college was found to result in academic failure, which contributed to an unacceptable drop-out rate for freshmen. Third, the assessment program identified the issue of minority retention as similar to, but at the same time somewhat more complex than, retention of majority students. By documenting retention rates, identifying variations in the rates among various groups, and identifying some of the underlying causes of this problem, the assessment program helped to focus the institution's attention on retention and to suggest possible courses of action.

In the case of the stress related to the challenge of adapting to the VMI culture, profiles were developed through the use of the Myers-Briggs Type Indicator (Briggs-Myers, 1976) of students who dropped out within the first five (rigorous) weeks of the semester. The conclusion drawn from studying these profiles and experimenting with various solutions to this problem was that these students could be assisted in their transition by "a greatly expanded orientation period with programs designed specifically to help students and their parents learn what to expect from VMI and how to overcome the stresses of the system" (*Third Assessment Report*, p. 62).

Regarding the stress that led to academic failure and attrition, it was found that "traditional measures such as Scholastic Aptitude Test scores have proven sporadic predictors of academic success at VMI" (*Third Assessment Report*, p. 63). Therefore, the Myers-Briggs Type Indicator and a learning thinking-styles inventory were used to develop profiles of at-risk students and to suggest possible ways to assist these students. The result has been a multifaceted approach including the establishment of a Learning Center, which provides a wide range of academic support; the hiring of Learning-Skills Specialists; the development of academic support modules; and the provision of individual counseling of students based on their learning profiles. In addition, an electronic database containing these individual student profiles was created to keep track of students and the changes that occur in their affective profiles, as well as changes in their academic performance. As part of a privately funded program, faculty were also provided with resources to make innovative changes in their courses to meet more effectively the diverse learning styles of students.

The collection and use of assessment information as part of the planning process has led to a synthesis of assessment and institutional life at VMI.

The issues related to minority retention were somewhat more complex, since there was a higher academic failure rate but lower attrition due to stress for minority students than for majority students. The profiles of minority students shed light on these differences, and specific programs were initiated for minority advising and academic support "as a means for assuring that minorities can take full advantage of the VMI experience" (*Third Assessment Report*, p. 67).

A number of other issues have been identified by the assessment program, which have been or are beginning to be addressed college-wide and in various departments and programs. It is clear from these efforts that in the last few years the collection and use of assessment information as part of the planning process has led to a synthesis of assessment and institutional life at VMI.

For more information about the assessment program at VMI, contact: R. Stephen RiCharde, Office of Institutional Analysis, Virginia Military Institute, Lexington, VA 24450-0304.

Assessment Update
March-April 1994
Volume 6, Number 2

References

Briggs-Myers, I. *Myers-Briggs Type Indicator*. Palo Alto, Calif.: Consulting Psychologists Press, 1976.

RiCharde, R. S., Olney, C. A., and Erwin, T. D., "Cognitive and Affective Measures of Student Development." In T. W. Banta and Associates, *Making a Difference: Outcomes of a Decade of Assessment in Higher Education*. San Francisco: Jossey-Bass, 1993.

Third Assessment Report, 1993. Lexington, Va.: Office of Institutional Analysis, Virginia Military Institute, 1993.

University of Colorado at Boulder

Institutional accountability is an ongoing outcomes-assessment process that focuses on activities that will help university units evaluate their programs to identify strengths and determine areas for improvement.

In January 1987 the chancellor of the University of Colorado at Boulder convened the Blue Ribbon Faculty Committee (BRFC) to develop a policy for assessing student outcomes. This action was in response to Colorado Revised Statutes 23-13-101, which established a higher education accountability program requiring the university by 1990 to institute a system to assess "demonstrable improvements in student knowledge, capacities, and skills between entrance and graduation."

The University of Colorado at Boulder, founded in 1876, has a total enrollment of 25,000 students, including 20,000 undergraduates, and approximately 60 academic programs at the baccalaureate level. The university is the only Rocky Mountain region institution in the American Association of Universities.

The Institutional Accountability Program of the University of Colorado at Boulder was developed within the context of the university's goals and objectives of the 1987-1988 strategic plan. The plan encouraged each department to examine critically the objectives of its curriculum and its contribution to general education requirements, to develop a process for the assessment of educational outcomes that would assist in the enhancement of curriculum and the improvement of teaching skills, and to enhance efforts to determine causes of student attrition. In addition, the BRFC report stated that institutional accountability is an ongoing outcomes-assessment process that focuses on activities that will help university units evaluate their programs to identify strengths and determine areas for improvement, help improve communication with the university's public constituents, and assist in reaccreditation reviews (Institutional Accountability Program, 1988).

From January to November 1988, departments, schools, and colleges developed their goal statements under the guidance of the Outcomes Steering Committee. These goals are now published as part of the university catalogue. Between November 1988 and May 1989, these units

Assessment Update
May-June 1994
Volume 6, Number 3

developed initial procedures to measure student attainment of their stated goals.

The BRFC guidelines specified the assessment of three areas: general education, education in the major discipline or specialty, and relevant activities that occur outside the classroom. General education outcomes were further defined to include the ability to gather and evaluate information; acquaintance with knowledge and concepts that provide a culturally appropriate matrix for the information and for decisions based on it; and the ability to communicate information, decisions, and the reasons for them to other people.

Very specific directions were given to units. They were asked to state explicit goals for undergraduates, to examine current programs in light of those goals, to choose (and/or develop) and implement ways to measure the achievement of those goals, to use the results of assessment to strengthen programs, and to use the results of assessment in regularly scheduled program and accreditation reviews. In assessing the achievements, outcomes, and gains of their undergraduates as a group, units were to build on existing procedures and add new ones as needed.

The BRFC guidelines specified the assessment of three areas: general education, education in the major discipline or specialty, and relevant activities that occur outside the classroom.

An October 1989 report includes the assessment plans for each of the academic units within each school or college (Institutional Accountability Program, 1989). This includes the plans of the 34 departments or programs in the College of Arts and Sciences, and those of the colleges or schools of business, education, engineering, environmental design, journalism and mass communication, and music.

Each unit in the College of Arts and Sciences, for example, was asked to prepare a plan to assess at least two goals, preferably one dealing with a knowledge goal and one with a skills goal. The College of Arts and Sciences general education program plan, which is embodied in the college's core curriculum revised for 1994, lists expected student outcomes in seven content areas: historical context, cultural and gender diversity, United States context, literature and the arts, natural sciences, contemporary societies, and ideals and values.

General education skill areas include written communication, foreign language, quantitative reasoning and mathematics, and critical thinking. The main student outcome in written communication, for example, is the ability to formulate and execute an extended written argument coherently, clearly, persuasively, and gracefully. The first student outcome in the critical thinking area is the ability to recognize and avoid common mistakes in reasoning.

Programs and departments within the college have their own sets of knowledge and skills outcomes, as do the other schools and colleges. The College of Engineering and Applied Science, for example, has knowledge outcomes in the areas of mathematics and basic science, engineering science, engineering design, and humanities and social sciences.

Assessment Update
May-June 1994
Volume 6, Number 3

Outcomes assessment procedures, results, and impacts are documented each year in a report prepared for submission to the Colorado Commission on Higher Education. First, the report documents changes in teaching, curricula, and learning made as a result of assessment. Second, it describes how the information is made available to the public, including students, and how students are informed of institutional expectations of assessment. And, third, it contains data on general education, discipline-specific education, retention and completion, alumni and student satisfaction, and after-graduation performance.

The report has been produced in three formats. The full report contains all campus data and summaries from each unit, including the goals assessed, methods used, validation, outcomes and evaluation, program changes, and future assessment plans and changes. The short-form format provides an overview of all-campus data and a summary table that indicates whether or not units have engaged in activities under various categories such as external review of goals, use of national tests, use of portfolio evaluation, proposed curriculum changes, and changes in assessment methods. The third format, an executive summary of the full report, briefly describes results related to student and alumni satisfaction. For example, 83% of respondents to a survey of 1993 seniors and 92% of respondents to a survey of 1989 graduates would recommend the University of Colorado at Boulder to a friend considering college (Institutional Accountability Program, 1993). Student retention and completion and after-graduation performance are additional elements of this third format.

Ninety-two percent of respondents to a survey of graduates would recommend the University of Colorado at Boulder to a friend considering college.

The executive summary also gives examples of specific findings. For example, assessment of the writing skills component of the College of Arts and Sciences general education program indicated that about 62% of all students tested showed recognizable improvement in their writing skills and 31% achieved a marked improvement. Also briefly summarized are changes resulting from assessment, such as the offering each semester, rather than just in the spring, of the senior seminar in the Department of Theater and Dance, with separate sections for theater majors and dance majors. Finally, the executive summary describes the procedures and procedural changes in assessment that have resulted from the previous years' assessment results. For example, the Department of Kinesiology plans to add an evaluation of students' skill in critically reviewing scientific research articles. These reports clearly indicate both the extensive effort that has gone into the development and implementation of the University of Colorado at Boulder Institutional Accountability Program and the impact that its implementation has had in the university.

The Office of Research and Information has provided professional help to all units in their efforts to prepare assessment techniques. For more information about the Office of Research and Information and the University of Colorado at Boulder Institutional Accountability Program, contact Ephraim I. Schechter, Senior Researcher, Office of Research and Information, University of Colorado at Boulder, Campus Box 108, Boulder, CO 80309-0108. Tel.: (303) 492-2516.

Assessment Update
May-June 1994
Volume 6, Number 3

References

Institutional Accountability Program. *5.00 Proposal of the Institutional Accountability Program*. Boulder: University of Colorado, 1988.

Institutional Accountability Program. *Report on the University of Colorado at Boulder, Academic Year 1988-1989*. Boulder: University of Colorado, 1989.

Institutional Accountability Program. *Report on the University of Colorado at Boulder, Academic Year 1992-1993*. Executive summary submitted to the Colorado Commission on Higher Education. Boulder: University of Colorado, 1993.

Miami University
Oxford, Ohio

Effective assessment makes use of a variety of methods and approaches, providing multiple measures on multiple groups over time.

Assessment Update
September-October 1994
Volume 6, Number 5

The assessment program at Miami University in Oxford, Ohio, has its foundation in a resolution passed by the University Senate in 1992 (SR92-A2). The resolution states that Miami University is committed to offering high-quality programs for its students and encouraging systematic collection of data to inform discussion about program effectiveness and improvement. Specifically, the purpose of assessment at Miami University is to foster thoughtful and informed reflection on the outcomes of academic and cocurriculum programs by relevant faculty, students, and staff within units. This purpose is different from that of program review information, which is intended for use in decision making by external audiences.

The resolution indicates that the following types of information should be collected to guide improvements in the learning environment at Miami University: incoming student characteristics, the nature and impact of educational experiences inside and outside the classroom, graduating student characteristics, and alumni evaluation of their experience. Assessment activities designed to gather this information should take place within individual programs, departments, colleges and schools, regional campuses, academic and student affairs divisions, and universitywide programs and services. These activities are linked to ongoing planning efforts and they are intended to provide useful information in response to questions of genuine interest and importance to each unit.

The questions and related assessments are tied to each unit's specific mission as well as to the contribution of that unit to the general university mission. In this regard, the resolution indicates that assessment information should not be used to make decisions about the performance of individual faculty members or students or reported in a way that compromises the anonymity of respondents.

The resolution recognizes that effective assessment makes use of a variety of methods and approaches, providing multiple measures on multiple groups over time. That is, assessment is viewed as an ongoing activity.

BEST COPY AVAILABLE

And sampling is suggested as a way to balance the costs and benefits of assessment over time.

The underlying assumptions of the Miami University assessment approach are framed as follows (Schilling, Schilling, Baxter Magolda, and Morenberg, 1993). First, it is noted that most of the work being done in higher education assessment has been evaluative in character, perhaps because the early call for assessment came from government officials asking about accountability, and because there is a general tendency in the United States to value comparative and competitive models. By this account, much of the effort that has gone into assessment thus far has resulted in frustration because it has not made a difference in the educational experiences of students on campus. In other words, since the evaluative efforts have provided summative rather than formative information, faculty do not get much information that will actually help them improve the quality of the curriculum or teaching on campus.

In their collection of materials on assessment, members of the Miami University Assessment Committee of the Liberal Education Council (Schilling, Schilling, Baxter Magolda, and Morenberg, 1993) suggested that in order for assessment to make a difference, it must be embedded in a framework that can help faculty and administrators understand and implement organizational change. Ideas related to Continuous Quality Improvement (CQI) are seen to provide a conceptual framework and a set of techniques that will foster the use of assessment results to improve the undergraduate experience.

Assessment is seen as a potentially powerful tool for making the curriculum visible, an important first step in seeking continuous improvement.

Within the context of CQI, the curriculum is described as an invisible system with lists of courses that are linked together, often by elaborate rhetoric, to create general education requirements, department majors, and degree requirements. Assessment is seen by those at Miami University as a potentially powerful tool for making the curriculum visible, an important first step in seeking continuous improvement as long as it does not make faculty defensive. In order to make the curriculum visible in ways that do not engender fear and suspicion among faculty, the Miami University assessment approach uses descriptive, rather than evaluative, techniques, such as comprehensive portfolios, student time-use studies, free writing, and intensive student interviews. It is believed that these approaches make the curriculum visible by emphasizing its intersection with the lives of students.

Beginning in 1990, several projects were initiated at Miami University to provide insights into the utility of descriptive assessment. One such experiment involves the use of portfolios by a random sample of 30 to 40 students from each entering class. These portfolios are intended to provide information on the nature of the liberal education program.

Comprehensive portfolios are described as collections of the artifacts of the curriculum and, as such, may include research papers and other process reports; multiple- and other forced-choice tests; self-evaluations, personal essays, and journals; computational exercises and problems; photographs

Assessment Update
September-October 1994
Volume 6, Number 5

of art projects; computer programs; case studies; creative writing; short-answer quizzes; and essay tests. These may be completed in class or as out-of-class assignments.

One analysis revealed that the typical portfolio contained 161 pages of materials, with personal essays making up 55% of the total and research papers constituting less than 8%. In addition, multiple-choice items were found to be the dominant mode of evaluation during students' first year, accounting for 52% of all examination questions. Narrative summaries of individual portfolios revealed the range of topics covered and diversity of students' experiences. The quantitative analyses and narrative summaries are used to generate tentative hypotheses regarding the meaning of materials.

Time-usage logs are kept by a select group of students for a seven-day period. Each student is given a watch that is set to beep at predetermined intervals. Students complete an activity log by writing the time that the beeper goes off and briefly (in four or five words) describing the activities that they are engaged in when beeped.

Faculty interview students annually throughout their careers at Miami University.

These logs are analyzed to determine the range and duration of student activities. For example, the analysis of one set of logs showed that approximately 36% of students' time was spent in academic activities such as being in class, studying, and meeting with advisers or tutors. Personal time (chores, errands, church, hygiene activities, walking, or driving) accounted for the next largest amount of time spent (approximately 24%). Students reported spending relatively less time engaged in athletics, socializing, leisure time, eating, sleeping, working, and so on.

Free writing activities are structured around sets of questions relevant to first-year, second-year, and third-year students participating in the portfolio groups. For example, first-year students were asked to respond to the following: "Before you came to Miami, what did you think college would be like? Write about the ways in which your Miami experience has changed or not changed your thinking about college." "What have you read, observed, heard, or done in the past semester that caused you to recognize and examine your assumptions about people different from yourself?" "Share what you liked best about your classes last semester and what you liked least." "Describe a course assignment that asked you to identify and work on a question, issue, or problem." Analyses resulted in lists that summarized the students' responses.

Faculty interview students annually throughout their careers at Miami University. Typically lasting from 45 minutes to 1 hour, these interviews are intended to provide follow-up information about the portfolio, time-usage study, and free writing exercise. During the interviews students are asked to reflect on what they learned from each activity and to elaborate on materials or ideas. Faculty at Miami University have reported gaining insight into how students experience courses, what they like and do not like about various instructional approaches, what is important about the

Assessment Update
September-October 1994
Volume 6, Number 5

classroom environment that facilitates or hinders their learning, and the nature of assignments that allow them to learn best.

Results of analyzing the portfolios, time-usage logs, free writing, and interview exercises have been shared with faculty members campuswide through periodic newsletters and with parents of incoming students. These results have made it possible to create a compelling description of students' education and experiences at Miami University. Faculty have used this information to take a close look at what is required of students, at the effectiveness of the curriculum, and at teaching, collectively. Parents are able to get a realistic sense of the work that students may be doing and the type of curriculum they will encounter. In essence, these activities have made the curriculum visible for discussion and improvement based on input from faculty, students, and parents.

For more information, please contact Karl L. Schilling, Western College Program, Miami University, Oxford, OH 45056. Tel.: (513) 529-1275.

Parents are able to get a realistic sense of the work that students may be doing and the type of curriculum they will encounter.

Reference

Schilling, K. L., Schilling, K. M., Baxter Magolda, M., and Morenberg, M. *A Collection of Materials on Assessment at Miami University*. Oxford, Ohio: Assessment Committee of the Liberal Education Council, Miami University, 1993.

Chicago State University

Faculty have complete responsibility for designing and implementing the CSU assessment plan.

Assessment Update
November-December 1994
Volume 6, Number 6

Assessment at Chicago State University (CSU) began in the fall of 1989 with the appointment of the Committee on Undergraduate Assessment. Formal assessment activities at CSU were initiated as a result of a mandate from the Illinois State Board of Higher Education. In addition, the Board of Governors of State Colleges and Universities regularly monitors the assessment process. And assessment activities are also reported through monthly Board of Governors reports, the undergraduate review, and the annual university report. The assistant provost for academic development chairs the committee and has primary responsibility for coordinating all assessment efforts and providing support and assistance to individual departments and faculty.

The assessment program at CSU focuses on improving learning outcomes. After careful consideration of what students are expected to have learned when they graduate from CSU, an assessment plan with the following components was developed by the committee. As described in the report by the committee (Lipscomb, 1993), assessment occurs within the context of the university's three-point model for student success, which includes precollege initiatives, program improvements, and safety nets and windows of opportunity for all students.

CSU has expectations pertaining to students' precollege skill levels in reading, mathematics, and composition. Thus, the first component of the CSU plan is the assessment of all entering students regarding these basic skill areas. All students earning the baccalaureate degree are expected upon graduation to be proficient in general education areas and in a specialty area or major. Those seeking a master's degree are expected to be proficient in their specialty areas at an advanced level. Therefore, the second component of the CSU plan includes the assessment of curriculum areas pertaining to undergraduate general education, the undergraduate major or specialty area, and the graduate specialty area. A large percentage of CSU students are academically underprepared minority students. Since CSU expects its students to complete their degree programs successfully, and is committed

BEST COPY AVAILABLE

to assisting them in this effort, the final component of the plan is an assessment of student support services.

The University Standing Committee on Assessment has as its purpose the facilitation of CSU's comprehensive assessment plan. Consistent with the mission of CSU, the assessment plan is designed to facilitate curriculum development and to encourage student learning and achievement, while serving as a vehicle for an integrative faculty development program and for student self-assessment. Faculty have complete responsibility for designing and implementing the CSU assessment plan, beginning with faculty representation on this universitywide committee. In addition, membership on each departmental subcommittee is composed of faculty who teach the courses being assessed, as well as other faculty in the department.

The faculty subcommittees began by identifying learning outcomes that were consistent with the university's mission, the recommendations of professional societies, the requirements of the State Board of Higher Education, and the guidelines of accreditation agencies. In developing learning outcomes, the subcommittees asked the following questions: What are the desired outcomes in terms of student behavior? What skills must students have? What human and professional values must they develop? What is being done to achieve these outcomes? How do specific experiences lead to the achievement of these outcomes?

The subcommittees asked the following questions: What are the desired outcomes in terms of student behavior? What skills must students have? What human and professional values must they develop?

Based on the answers to these questions, departmental faculty on the subcommittees participated in the development or selection of appropriate assessment methods and their pilot-testing, administered instruments and other methods, analyzed and reported their findings, and used the results to improve their academic courses and programs. This involvement ensured faculty ownership of the departmental assessment process.

The use of multiple methods is an essential part of the CSU assessment plan. As noted in the Guiding Principles for Assessment at CSU, "Different methods of assessment may be appropriate for different programs. The methods of assessment of each program may include, but are not limited to, the following: portfolios, research projects, interviews, oral and written examinations, external evaluators, internships, senior seminars, senior recitals, exhibits and projects, and the practicum" (Lipscomb, 1993, p. 3). In addition, the following principle is stated: "Assessment strategies adopted by the faculty should be based upon clear specification of learning objectives for each course." The principle also states: "Assessment within the major should stress the development of skills related to the major, the development of human and professional values, and the knowledge of subject matter. The relationship among these three objectives should be established by each program's faculty.

"Each course in the curriculum should have well-defined objectives, including a specification of the student behaviors or performances sufficient to satisfy the objective.

"Multiple-choice examinations, if used, should be combined with additional assessment tools, such as written, oral, or behavioral exercises.

"Faculty development should include workshops on the techniques of choosing and/or developing reliable and valid instruments" (Lipscomb, 1993, p. 3).

As a result of adherence to these guiding principles, feedback has been provided to faculty and students. For example, the basic skills assessment, which is more than fifteen years old, provides feedback to faculty and students regarding reading, English, and mathematics. This feedback is based on a diagnostic placement examination taken within the first semester after a student is admitted and on a second version that must be taken after the completion of the courses. As a result of assessment activities during 1990-1994, several changes have occurred in these developmental programs, including the addition of a fourth hour of instruction, increased faculty training and increased coordination in mathematics and English, a revision of the developmental program in English composition, and development of ancillary teaching material in mathematics (for details, see Lipscomb and Billimoria, 1994, pp. 13-14).

Faculty in seventeen academic departments have used assessment information to make changes since 1990.

In most departments the assessment plans related to majors include some kind of integrative experience and at least one other measure of student learning. Faculty in seventeen academic departments have used assessment information to make changes since 1990. For example, art options have been totally revised and now require art exhibits at the end of the sophomore, junior, and senior years. Similarly, courses have been added to the English education major and to other College of Education majors. Beyond the assessment procedures in the major, CSU also has worked to clarify and strengthen integrative components in baccalaureate programs (for details, see Lipscomb and Billimoria, 1994, pp. 14-16).

Students are given feedback regarding their performance on basic skills assessments as part of developmental courses. Results are discussed with students so that they can make informed decisions about the steps needed for success and what they can do to aid their own learning.

Academic support programs follow the same assessment steps as the academic departments. That is, first, outcomes statements were developed to guide the selection of appropriate assessment methods, and then the assessment process was implemented using those methods to determine the extent to which outcomes were being achieved. Assessments used by academic support programs include student satisfaction surveys, attitudinal and opinion surveys (that is, the American College Test [ACT] Entering Student Survey, the ACT Withdrawing Student Survey, and the ACT Opinion Survey for graduating seniors), volume of activity logs, and level of proficiency counts related to the initial and follow-up basic skills placement examinations. As of the end of the third year of the assessment of student support programs, assessment data have been used to improve the early warning system, advising, student support services, and counseling.

Assessment Update
November-December 1994
Volume 6, Number 6

Improvements in the assessment program itself have occurred as a result of feedback, since the committee conducts an annual review of the university's assessment activities. During the first review, outcomes for each program were assessed to determine their feasibility for instrument development. Feedback in this regard was given to departmental chairs and the assessment subcommittees. During the second and third years, departments' activities were monitored so that they could be moved toward data collection and analysis. In the fourth year, assessment coordinators were appointed for each academic department, and a system for monitoring the use of assessment findings to facilitate curriculum and program modifications was implemented. Still to be accomplished are, first, the dissemination of findings to all areas of the university, and, second, the assimilation of general education and graduate programs into the assessment process.

For more information about the CSU assessment program, please contact Delores Lipscomb, Assistant Provost for Academic Development, Cook Administration Building, Room 129, Chicago State University, Chicago, IL 45056. Tel.: (312) 995-3860.

Improvements in the assessment program itself have occurred as a result of an annual review of the university's assessment activities.

References

Lipscomb, D. *Assessment at Chicago State University: Improving Learning Outcomes.* Chicago: Chicago State University, 1993.

Lipscomb, D., and Billimoria, A. *Review of Undergraduate Education, 93-94.* Report to the Board of Governors of State Colleges and Universities. Chicago: Chicago State University, 1994.

Western Carolina University

The role and mission statement of the university asserts that "the most important activity is student-teacher involvement in learning."

Western Carolina University (WCU) is a regional, comprehensive university with an enrollment of 6,400 students. It attained university status in 1972 with the offering of graduate programs beyond the master's degree. Approximately 60% of enrolling students are first-generation college students who bring a wide range of academic ability and preparation.

Faculty at WCU view teaching as their major commitment. As a result, the role and mission statement of the university asserts that "the most important activity is student-teacher involvement in learning." In addition, the fundamental role of the university is "to develop a community of scholarship in which students, faculty members, administrators, and staff members learn and apply the products of learning" (Stillion, Hinson, and Corbin, 1993).

WCU's assessment program began in 1987 with the appointment of the Outcomes Assessment Advisory Council. Composed of faculty and administrators, the council worked for three years to develop the principles and guidelines for assessment at WCU. The result of this work was an assessment resource guide (see Corbin, 1994), which was first distributed to all departments in 1989. The following basic principles guide assessment efforts at WCU.

1. The primary aim of assessment is to involve faculty in the evaluation of academic programs in order to provide evidence of program quality, permit documentation of changes in program quality, establish a longitudinal data base documenting consistently high academic standards, and show evidence of student learning.

2. The WCU assessment program is designed to measure the skills, knowledge, and accomplishments specified in the role and mission statement.

3. Assessment is an ongoing process, designed by the faculty in keeping with the role and mission statement of the university and of the schools and departments in which programs are located.

Assessment Update
March-April 1995
Volume 7, Number 2

4. Defining goals for each program area is the most important step in establishing an assessment program. In each academic major, goals are defined, students' progress toward goals is monitored, and results are used to continue, modify, or reinforce aspects of the program to ensure continuous advancement in program quality.

5. No single approach to assessment is universally accepted; therefore, different methods are appropriate for different programs. The faculty for each program are responsible for determining the methods and appropriateness of the assessment measures used. (Four types of methods are described.)

6. Although students' performance on assessment measures will be used to evaluate programs, scores on assessment measures may not by themselves be used to impede student progress toward graduation.

7. Assessment will focus on programs rather than individual faculty performance.

8. Program evaluation measures will be considered confidential. Identities of individual people will not be released.

No single approach to assessment is universally accepted.

The WCU undergraduate assessment process involves all areas of the curriculum, including general education, academic majors, and the honors program. The development and implementation of the WCU undergraduate assessment plan are consistent with the North Carolina legislative mandate that requires "each institution to develop a plan that would exhibit how the institution will measure its effectiveness, especially in the areas of student learning and development, faculty development and quality, and progress toward the institution's mission" (Faculty Handbook, 1992, p. 23).

The General Education Committee is responsible for assessing the general education program, which consists of 41 semester hours and is divided into two areas: 16 hours of foundations and 25 hours of perspectives. Foundation categories include written communication, mathematics, oral communication, computer literacy, and leisure and fitness. The perspectives area includes human behavior, contemporary institutions, physical and biological sciences, the humanities, fine and performing arts, comparative cultures, and the human past.

As described by the Office of Assessment (1994), 12 focus groups composed of faculty who taught the actual courses in these areas were formed in 1989 to discuss course content, teaching methods, and shared problems. In the second year of operation, each focus group developed educational goals for students taking a course or courses in its area. These goals represented the common minimum core of knowledge, skills, attitudes, and understanding that students should have regardless of which course they took in the area. Next, the focus groups began formulating and testing assessment strategies for measuring progress toward attainment of their stated goals.

The use of focus groups has placed the faculty at the heart of assessment of students' accomplishment of educational goals within each area of general education. For example, within the foundations and perspectives areas are seven skills: written communication, oral communication, scientific methods, logical reasoning, interpretation, understanding of the valuing process, and reference and resource skills. With the exceptions of reference and resource skills and understanding of the valuing process, skills are measured among a sample of freshmen using the Assessment of Reasoning and Communication from the American College Test. A sample of seniors has also been assessed. The assessment will be repeated with the same group of students when they have completed between 45 and 60 semester hours. The collection of longitudinal, national data permits both external and internal comparisons.

The use of focus groups has placed the faculty at the heart of assessment of students' accomplishment of educational goals.

The assessment of general education goals has several other components in addition to these skill-related student assessments. Student and faculty surveys constitute another component. These surveys are given in two categories of general education each semester. All categories are surveyed on a three-year cycle as a part of program review. The surveys focus on course content, course syllabi, handouts, tests, and examinations.

The final general education assessment component includes various universitywide surveys. For example, incoming students (both first-year and transfers) and graduating students are surveyed about their perceived levels of preparation on each general education skill objective, as well as on each foundation and perspective area. In conjunction with each major, a series of alumni surveys have been developed to determine how well graduates feel they are prepared. Employers also are contacted (with the permission of alumni) to gather independent perspectives on the preparation of graduates.

Assessment in the major is based on plans developed by the faculty in those majors. As with the general education assessment, departments report the data for each major according to the major assessment plan's goals and instruments.

All of the information collected about the general education program and the majors is entered into the university's longitudinal assessment data base. As a result, detailed questions can be answered about what changes need to occur, what changes have occurred, and the effect of those changes on students' learning and perceptions.

The initial implementation of the WCU assessment program has resulted in three types of changes (Office of Assessment, 1994). The first involves specific changes in courses, programs, and majors, such as offering more sections or more regular scheduling of courses, making sure syllabi are available in all courses, addressing class absence policies, and ensuring that general education skills are incorporated into the appropriate courses.

Assessment Update
March-April 1995
Volume 7, Number 2

The second type of change has been the development of more specific information for both faculty and students, such as the definitions of the general education seven skills areas, which have been accepted by the Faculty Senate and the chancellor. These statements not only permit the evaluation of student learning but also provide a common vision for student achievement.

The third type of change was fostered by the discovery that there was less coherence in the curriculum than had been assumed. As a result, for example, the thinking, reasoning, and expressing category was replaced by an oral communication requirement, since the faculty focus group responsible for this category could not agree on common educational goals for the current courses.

From the point of view of faculty at WCU, the process that they used to develop and implement their assessment program worked for several reasons (Office of Assessment, 1994). First, a director of general education was appointed at the request of the faculty. Second, a representative committee of faculty meets almost weekly to monitor the evaluation process. The General Education Committee is responsible for ensuring that the program provides a general base for all majors. Third, focus groups, composed of faculty who teach the general education courses, created both the educational goals and the methods for assessing them. Therefore, there is faculty ownership in the process. Finally, support and cooperation are provided by four offices that are responsible for implementing the assessment and program review processes. These include the Office for Academic Affairs through the associate vice chancellor, the Office of Assessment, the Office of Institutional Studies and Planning, and the Office of General Education.

Support and cooperation are provided by four offices that are responsible for implementing the assessment and program review processes.

For more information about the assessment program, please contact Judith Stillion, Associate Vice Chancellor, Academic Affairs, Western Carolina University, Cullowhee, NC 28723-9015. Tel.: (704) 227-7495.

References

Corbin, R. P. *Assessment: A Resource Guide*. (2nd ed.) Cullowhee, N.C.: Western Carolina University, 1994.

Faculty Handbook, Vol. 3: Curriculum Development and Revision. Cullowhee, N.C.: Western Carolina University, 1992.

Office of Assessment. *Assessment of the General Education Program*. Cullowhee, N.C.: Western Carolina University, 1994.

Stillion, J. M., Hinson, F. D., and Corbin, R. P. "Assessment of General Education at Western Carolina University: Faculty-Owned, Continuous Improvement Approach." Paper presented at the American Association for Higher Education Assessment Forum, Chicago, 1993.

Sinclair Community College

At Sinclair, assessment is the means to the end of increasing institutional effectiveness.

In 1988, Sinclair Community College in Dayton, Ohio, formed the Steering Committee on the Assessment of Student Learning and Development to address the issues related to student outcomes, assessment, retention, and institutional effectiveness that are facing community colleges across the nation. The formation of this committee followed the development of learning outcomes, which began in 1986. In *Assessment of Student Learning and Development: Progress Report Focusing on Assessment*, Sinclair Community College (1995) provides both the background and current status of its effort to develop the Sinclair model for student outcomes assessment.

From the common set of issues facing community colleges, two conclusions were drawn by the steering committee: assessment tries to determine what students actually achieve in their college study, and assessment links educational objectives to measures of student outcomes. Assessment also was seen to require a process of self-examination, which should be motivated by one central goal: to improve student learning. There are three steps in this process of self-examination: articulation of the goals for each academic program, evaluation of students' progress toward those goals, and use of the resulting information to continue, modify, or reinforce aspects of each academic program to ensure students' achievement. In summary, at Sinclair, assessment is the means to the end of increasing institutional effectiveness.

Since the clear emphasis of the higher education assessment movement, and assessment at Sinclair, is on the measurement of student progress toward institutional goals and objectives, a comprehensive program of assessment was developed at Sinclair that begins before the time of enrollment, continues through enrollment, includes outcomes assessment at graduation, and then continues for some reasonable time after graduation. In the Sinclair model for student outcomes assessment, three sets of assessment indicators have been or are in the process of being created to provide information about student progress. Because there are advantages

Assessment Update
May-June 1995
Volume 7, Number 3

and disadvantages to different assessment approaches and methods, the Sinclair model combines a variety of methods and target populations to create an efficient and cost-effective program.

One set of assessment indicators is intended to gather information on all degree-seeking students. This set contains seven indicators: student goal attainment from preenrollment to postgraduation; preenrollment, entry-level basic skills; academic programs learning outcomes during enrollment; general education learning outcomes after completion of 62-75 quarter hours; graduate satisfaction at the point of graduation, one year after graduation, and three years after graduation; placement after graduation for those with A.A.S. degrees; and transfer success after graduation. A second set of assessment indicators is intended to gather information periodically from a sample of select student populations. This set includes student goal attainment in relation to academic departments and student services; graduation rates; satisfaction of current students, former students, and graduates; tracking of developmental studies students' retention and success; and licensure or certification outcomes, when appropriate. The third set of assessment indicators includes employer satisfaction surveys such as those that academic departments typically use in response to accreditation requirements.

The progress report notes that "as an open door urban community college, Sinclair is committed to providing access coupled with high quality instruction to all interested community members. . . . As a successful community college, Sinclair must be open to students of all types, and must create an environment conducive to student success once the student enters the college. Assessment, especially entry level assessment, is a crucial factor in providing individuals with opportunities to enter college and for succeeding once enrolled" (Sinclair Community College, 1995, p. 28). The placement testing, diagnostic testing, and developmental studies courses are all part of this commitment to students' success.

Entry level assessment, is a crucial factor in providing individuals with opportunities to enter college and for succeeding once enrolled."

In October 1993, Sinclair began using the Computer Adaptive Placement Test (COMPASS) developed by the American College Test company. COMPASS is a seamless battery of assessment instruments designed specifically for use by community, technical, and junior colleges. It consists of assessments of reading skill; language usage; numerical skill; and elementary, intermediate, and college algebra. The results are given to students and to their academic counselors and faculty advisers. Assessment results are used to place students correctly, including placement in developmental studies courses, if necessary.

The Steering Committee on the Assessment of Student Learning and Development made three recommendations regarding learning outcomes assessment: the assessment of skills in general education and in the major should be mandatory for all degree-seeking students, but graduation should not be contingent on the testing results; higher-order skills such as critical thinking and knowledge and skills in a student's program of study should also be assessed as deemed appropriate by the department; and there should be neither logistical nor monetary barriers to a student's participation in learning outcomes assessment.

All degree- or certificate-seeking students must be assessed, as is the case with similar transfer students who do not have transferable equivalent mathematics or English courses. In fact, degree- or certificate-seeking students cannot register for any credit course until they have test scores on file. Some programs at Sinclair also use specially focused diagnostic assessments such as tests for mechanical reasoning, writing, health aptitude, and other professional skills.

Consistent with these recommendations, the purpose of assessing learning outcomes in general education and in academic programs is to determine whether growth and changes have occurred throughout the student's college experience. Minimum competence in basic skills and knowledge and in higher-order skills and knowledge is assessed. Assessment strategies are designed to measure effectively those skills and knowledge areas critical to a new graduate's successful entry into the work force or into parallel university programs. They are not intended to be used punitively against students, and the assessment of student learning and development is a process that is separate from faculty evaluation.

The purpose of assessing learning outcomes in general education and in academic programs is to determine whether growth and changes have occurred throughout the student's college experience.

Academic departments at Sinclair have many different options for assessing student learning: commercially available, standardized national examinations; comprehensive examinations developed by peers, faculty, or external evaluators; task and talent performances as part of field experiences (for example, internships and student teaching); simulated task performances such as "in-basket" projects, simulated work experiences, and computer simulations; moral and ethical choice exercises; concept mapping and Gowin's V-diagram development by students; holistic assessment of comprehensive research projects, papers, theses, portfolios, performances (videotaped), or exhibits involving an oral defense or interview; interviews structured by a discipline-specific conceptual map; open-ended individual or group interviews; group problem-solving tasks; programmatic growth contracts; proficiency checks of psychomotor skills; course-embedded assessments as part of seminars or capstone courses; surveys of current and former students involving self-reports of knowledge gained, changes in skill levels, and levels of involvement in the learning process; and tracking of student groups over time to determine persistence, program progress, program completion and time to completion, course passage rates, and grade point performance.

In addition to the general education and department-specific assessments, there are a number of collegewide student-tracking projects. These involve periodic sampling of select populations. For example, an employer survey is conducted to determine, from the employers' point of view, the areas in which Sinclair graduates are best prepared and the areas in need of improvement. Technical education placement and academic transfer studies also are conducted. These are in addition to surveys of current students, former students, and graduates.

Assessment Update
May-June 1995
Volume 7, Number 3

Sinclair also is one of five institutions participating in a pilot project to implement a statewide student-tracking program. The tracking projects in this program include patterns of progress among developmental education students (a three-year longitudinal analysis of developmental education impact); program major progress and retention; descriptive reporting of characteristics associated with success among graduates; student goal certainty indicators; graduation rates and analysis of leavers among cohorts of first-time, full-time, degree-seeking students.

To learn more about the Sinclair model for student outcomes assessment and its contribution to academic introspection, recruitment and planning, readiness for accreditation studies, and improvements in teaching and learning at Sinclair Community College, please contact Linda Denney, Assessment Coordinator, Sinclair Community College, 444 West Third Street, Dayton, OH 45402-1460. Tel.: (513) 226-7961.

Reference

Sinclair Community College. *Assessment of Student Learning and Development: Progress Report Focusing on Assessment.* Dayton, Ohio: Sinclair Community College, 1995.

Sinclair also is one of five institutions participating in a pilot project to implement a statewide student-tracking program.

Assessment Update
May-June 1995
Volume 7, Number 3

St. Cloud State University

The faculty played a controlling role in developing assessment strategies and goals.

Guest Columnists: Philip Keith, Sandra Keith, Linda Lamwers, Karen Schmid.

St. Cloud State University has approached the assessment experience under several interesting conditions. First, we were initially pushed by our accreditation agency (North Central) into using assessment to rationalize our general education program. Second, because our governance structure is based on collective bargaining between the faculty and administration, the faculty has played a controlling role in developing assessment strategies and goals. We have developed a basic philosophy of assessment and a plan that includes but is not limited to assessing academic achievement and student learning. In getting this far, we have constructed instruments for general education and other academic programs that are both homegrown and shaped by visits by various assessment consultants over the past several years—Peter Ewell, Pat Hutchings, Mary Ann Bunda, John Centra, Parker Palmer, Michael Knight, Donald Lumsden, and Anita Fellman. A number of our faculty and administrators also have had the opportunity to hear and interact with Patricia Cross, Tom Angelo, Alexander Astin, Jerry Gaff, and T. Dary Erwin; and over the years such experiences have given the process something like a glacial inevitability that has been both a strength and a problem.

In general education, we have a somewhat conventional "cafeteria" program with global program objectives and specific requirements in English composition and speech; distribution requirements in philosophy, social science, natural science, and arts and humanities; and a multicultural, gender, and minority requirement. The 1987 visiting team from North Central observed that (1) although we had a fairly rigorous process for approving general education courses within this framework, we had no way of monitoring their continued effectiveness in meeting general education goals; and (2) general education curriculum decisions seemed more a function of departmental convenience than of a philosophy of learning.

Assessment Update
September-October 1995
Volume 7, Number 5

As a collective bargaining institution, we operate more by consensus than by administrative design, and this fact tends to make radical restructuring unlikely. In response to the North Central concern about general education, we established a faculty-administration task force to look at the assessment issue. This committee recommended approaching assessment as a way of improving rather than summatively evaluating our operations and making assessment congruent with the existing process insofar as possible. Thus, assessment procedures were to reflect and improve existing curriculum and program review processes. Such an approach, we hoped, would draw the faculty at large into assessment because of their self-interest.

To start the general education assessment process, we requested that all general education courses be redescribed to document how they meet current general education program objectives. We also crafted a survey to tap student perceptions of the extent to which these objectives operated in the courses. Finally, we initiated a revalidation process under which all general education courses undergo a review every five years by a faculty general education committee and the Office of Academic Affairs.

Transforming a course approval mechanism into a course assessment mechanism is a conservative strategy in the sense that it does not encourage radical program revision; it is radical in forcing us to take the objectives much more seriously. Assessment has thus become the basis for examining and shaping our program philosophy; we can use it heuristically to help us identify problems in individual courses and in the program as a whole.

Transforming a course approval mechanism into a course assessment mechanism is a conservative strategy in the sense that it does not encourage radical program revision.

Some Illustrations

Two illustrations can show how this process works. Our general education courses must meet at least three of the following five objectives: (1) to provide an opportunity for work on basic academic skills-writing, reading, mathematics, and oral expression; (2) to provide an opportunity to see how disciplines are interrelated; (3) to provide an opportunity for work in critical thinking; (4) to incorporate a values orientation; and (5) to incorporate perspectives outside the student's experience with reference to culture, gender, and majority-minority status.

Operationalization of the first objective has already been useful in addressing the role of mathematics in general education. When we compiled the results of the first administration of the student perception survey, we found that mathematics barely registered in student perceptions of general education course content. Only 9 courses out of the 160 surveyed produced mean scores of 3.5 or above on a scale of 1 to 5. That made it rather clear that mathematics was not playing a very large role in the general education curriculum. If this institution is aiming to provide experience in mathematics as a "basic" function for all students, we are lacking; and in the current climate of concern about college students' ability to work in an increasingly technological society, the discovery was uncomfortable.

We knew that because the general education program did not encourage course prerequisites, many science courses had been largely demathematized. For example, mathematics is not a requirement; rather, it is one of several options within the natural science distribution area. However, because we were looking at a single measure of mathematics content, we had to triangulate with other data. When we reviewed a selection of graduating student transcripts, we found that about 75% actually had taken at least one mathematics course. Basic courses such as the business college's mathematics prerequisite and the calculus courses were not general education courses, therefore our program analysis was not "catching" them. Furthermore, since the objective was actually phrased in terms of mathematics calculation rather than quantitative reasoning or some other, broader equivalent, it seemed that many students were not seeing such activities as working with numerical tables and graphs, or reading numerical analyses, as fitting the survey criterion.

Our programmatic emphasis on interdisciplinarity in general education has been the source of some pride.

The issue of mathematics in general education has thus been brought into focus on three levels: as a curriculum-design issue, as an issue of how we describe and define the objective, and as a course-design issue. At the June 1994 American Association for Higher Education Assessment and Continuous Quality Improvement conference in Washington, Alexander Astin observed that standardized test scores suggest that on the average the mathematics ability of students declines in college. We share this view and believe that the pattern is a problem that we will need to address at St. Cloud by clarifying our objectives, establishing a quantitative thinking emphasis or requirement, and making faculty and students more sensitive to both the use and importance of mathematical reasoning activities.

Already in our course revalidation process we see several departments, including history, documenting the use of tables and numerical data, and we expect that along with this documentation will come increased attention to the nature, value, and misuse of numerical and quantitative information. In the Mathematics Department, the general education mathematics course, Cultural Math, has become a subject of interest. Defined in the catalogue simply as "topics chosen by the instructor to demonstrate the nature of mathematics," the course frequently has included extended units on primitive and ancient numerical systems and avoided work involving algebra, since that would place some students at a significant disadvantage given the difference in our students' mathematics backgrounds. We anticipate that the assessment-generated emphasis on the operationalization of program objectives and on the need to document continuity across different sections of the same course will continue to stimulate constructive review.

Assessment Update
September-October 1995
Volume 7, Number 5

The second illustration relates to the second objective—helping students see the relationships among disciplines. Our programmatic emphasis on interdisciplinarity in general education has been the source of some pride. However, the student survey showed that the students were largely unaware of this effort, in contrast to their awareness of widespread efforts to emphasize values and multicultural perspectives. This finding raises interesting questions: To what extent can the objective be meaningful if the students are unaware of it? How can we increase student awareness of this admittedly abstract issue? What are the effects of imposing this objective on courses? Predictably, we have found that when faculty members look at the low mean survey scores for this item, they tend to back away from claiming that their courses meet this objective, and the result sharpens the question of the meaningfulness of the objective.

On the whole, these are energizing questions. Is biology intrinsically interdisciplinary because of its reliance on chemistry and mathematics? To make a physics course interdisciplinary, does it need to reach for engineering applications in novels by Thomas Pynchon? Such questions have become the focus of discussion and class experimentation and research in several different areas, and they promise to generate even more attention as more general education courses proceed through the revalidation process in the next three years.

Validation, Review, and Reporting

There has been some worry about the current lack of objective validity in our surveying and reporting instruments, that is, about "the emperor having no clothes." This is a significant concern, given North Central's requirement that institutions be able to show that they are using learning outcomes data for program improvement. We are requiring learning outcomes data in the second cycle of the revalidation process and are encouraging and showcasing experiments with such data when used in the first cycle. Furthermore, we are developing projects for assessing learning outcomes in communication and critical thinking at a global program level using portfolios to meet North Central's new requirement.

Institutions must be able to show that they are using learning outcomes data for program improvement.

Our efforts to assess major, minor, and licensure programs, as stated above, take account of other assessment expectations that currently operate under external review and accreditation review programs. To document program improvement, the university assessment steering committee requested assessment plans of all degree-granting programs. To facilitate this process, the committee designed a reporting form that includes specific questions about program objectives and structure, description of students, assessment focus and needs, and a matrix incorporating institutional mission objectives and basic assessment techniques. This common reporting format will make it possible to develop an institutional picture of how the curriculum supports mission goals, and of the institution's use of assessment methods. Over time, the format will help to document the institution's coming of age in the methodology of assessment and to support enriched networking among faculty and programs.

How the reporting process generates engagement in assessment is illustrated in the Mathematics Department. The steering committee's request for assessment plans created fairly widespread concern and frustration this past spring, in spite of several workshop programs designed to facilitate the process. When the B.A. committee of the Mathematics Department sat down to address this request, their initial consensus was that the program could document its effectiveness in maintaining tough grading standards with the low grade point average for mathematics courses. They did feel somewhat frustrated, however, that this was all they could do, so they began to think about exactly what outcomes the program was emphasizing. They wanted to argue that beyond specialized knowledge of mathematics and scientific and quantitative thinking the program also addressed liberal knowledge and critical thinking. Course grades could be used to document these outcomes, and evaluation by practicum advisers, where possible, could be useful as well. To supplement these, they began to think that surveying of students, graduates, or both could be useful. With this broader sense of instrumentation, other outcomes began to seem worth thinking about: professional preparation and attitude changes, for instance. The Mathematics Department has in recent years been frustrated by a decrease of student interest in the major, and members could see how assessment considerations might actually become useful in helping them attract more students by developing and recruiting for the math minor.

They began to think that surveying of students, graduates, or both could be useful.

For more information about the assessment program at St. Cloud State University, please contact Philip Keith, Assessment Coordinator, 106 Riverview, 720 Fourth Avenue South, St. Cloud, MN 56301.

Assessment Update
September-October 1995
Volume 7, Number 5

PETER J. GRAY

Indiana University Bloomington

Indiana University Bloomington (IUB) is a large public Research I University with some 26,000 undergraduates, 7,600 graduate students, and 1,400 tenure-line faculty. The campus is organized into 12 schools, 53 departments, and 313 programs. Bloomington is a residential campus, and more than 90% of the undergraduate students enroll on a full-time basis (see Campus Assessment Plan, 1995).

Assessment was added to the mission statement.

In 1995 IUB celebrated its 150th anniversary. Throughout its history, a guiding principle has been that its faculty must be composed of scholars who are effective teachers. With the advent of the assessment movement, the university began to recognize how more systematic and comprehensive assessment of student learning could contribute to the traditional liberal arts values and goals underlying its mission.

Therefore, assessment was added to the mission statement as follows: the campus is committed to a high degree of faculty involvement in undergraduate education. It continues to review the curricula, use educational technologies, assess student learning, and provide coherent learning experiences and effective instruction. This mission statement implies that assessment is not to be added to instruction but rather must be part of it. It is assumed that assessment will build on current instructional and evaluation activities, will be consistent with the mission statement, and will be carried out by faculty in the course of their instructional duties.

In an act that further underscored the linkage between teaching and assessment, the deans of IUB's 12 schools drafted and approved a statement of assessment principles, which established basic conceptual and procedural parameters for assessment plans at the university, school, and department levels. Since the principles were in response to faculty questions about whether assessment was concerned with improvement or accountability, and how assessment outcomes would affect individuals and departments, they have helped faculty better understand institutional expectations. In general, all schools agreed that (1) faculty will participate in all phases of

Assessment Update
November-December 1995
Volume 7, Number 6

the assessment process, (2) student learning will be assessed using multiple measures, and (3) assessment data will be used to enhance instruction and curriculum.

The principles, in particular, emphasize improvement in instruction and student academic performance as the goals of assessment, with academic programs rather than individuals as the primary units of analysis. Wolf (1993, p. 4), in the Assessment Handbook, defined assessment as "research and inquiry into the improvement of teaching and learning."

The objectives that guide assessment at IUB are those learning outcomes that units concerned with general education, undergraduate majors, and graduate education seek to achieve. These objectives were developed from existing statements, and therefore they received ready approval from administrators and faculty. They also have the substantial advantage of relating "new" assessment initiatives to academic skills, knowledge, and values that are widely known and endorsed by faculty.

Each academic unit was responsible for defining objectives and developing assessment plans that evaluate whether department curriculum and pedagogy achieve student learning outcomes as defined. While many of the planned assessment activities were drawn from or built directly on existing faculty and administrative efforts, it was clear that more could be done. That is, improvements were sought where learning objectives were less clear than they should have been or where present indicators failed to provide sufficient information about students' knowledge, abilities, and attitudes.

The overarching principles and the objectives articulated at each level help ensure a close linkage between the core mission and values of the institution and the specific assessment initiatives implemented by the campus and its units.

The overarching principles and the objectives articulated at each level (general education, the major, and graduate education) help ensure a close linkage between the core mission and values of the institution and the specific assessment initiatives implemented by the campus and its units. An example of this linkage is the assessment of the general education program. Eight general education objectives were derived by the Office of Academic Affairs and an advisory committee of faculty from the course bulletins published by the academic schools. The objectives were developed to represent the learning outcomes that all units concerned with undergraduate education seek to achieve and were intended to provide a good summary of the cognitive skills students should acquire in a successful general education program.

The centerpiece of activities aimed at assessment of general education is a newly developed test of critical thinking, the Indiana University Student Performance Measure (Farr and Greene, 1994). This authentic performance measure requires students to read approximately one to two hours of materials on a given topic and to respond in essay format to open-ended questions on the readings. Students are rated on a holistic 4-point scale using established criteria for determining problem-solving and analytical reasoning ability. Pilot work has shown that the measure can be successfully administered to large numbers of students and that students and faculty value the exercise as academic work.

Assessment Update
November-December 1995
Volume 7, Number 6

A unique feature of IUB is its responsibility-centered management (RCM), which provides the management structure for assessment. RCM charges academic units with fiscal and academic planning, decision-making, and budgetary authority. In RCM, all revenue is attributed to the unit that generates the revenue. This structure invests the resource and planning functions essential to assessment in the faculty and unit administrators who must carry them out. The RCM structure combines faculty's traditional academic responsibility with increased control over budget and management, thereby promoting a closer alignment of academic priorities and institutional activities.

Because RCM is a fundamentally decentralized organizational structure, no initiative can succeed without support from local faculty and unit administrators. It encourages assessment of student outcomes, first, by providing academic and financial incentives to improve teaching and learning; second, by allocating resources at the appropriate level through administrative decisions that are local or unit-based; and, third, by making faculty committees and task forces integral to the decision-making and planning process.

Because RCM is a fundamentally decentralized organizational structure, no initiative can succeed without support from local faculty and unit administrators.

At all levels (general education, the major, and graduate education), faculty committees establish learning objectives for students, select and implement assessment measures, and act on assessment data collected. Central administrative and faculty governance bodies coordinate and disseminate information among units. Beyond the establishment of the guiding principles, the deans are actively engaged in the review of plans and of emerging assessment data. The chancellor appointed the vice chancellor for academic affairs to oversee development of the institutional plan and to work with the deans on broad issues of timing, policy, and direction. A full-time director of assessment meets with the associate dean of each school to coordinate university, school, and department plans. Each associate dean heads his or her school's assessment committee and, in turn, serves on the Campuswide Assessment Committee through which information, feedback, and approval flows. The associate deans report jointly to their own school deans and to the vice chancellor for academic affairs. The director of assessment in the Office of Academic Affairs works with departments to identify or design assessment measures and to aid in the analysis and interpretation of findings. In addition, there are some central "seed" funds available for assessment.

IUB admits a wider range of students than many other comparable Big Ten schools and, as a result, experiences a tension between maintenance of high academic standards and access. Faculty and administrators take seriously their responsibility to do all that is possible to enable all admitted students to achieve university standards and to graduate. Therefore, retention and assessment are closely connected activities in that assessment can be used to diagnose student needs and instructional shortfalls, to accommodate this wide range of students, and to improve retention.

Assessment results have prompted change at IUB. For example, a new writing test administered during orientation for first-year students was used to identify individuals with especially serious deficiencies. An analysis of advising and registration procedures indicated that course placement did not always accommodate these students. In response, the university developed a new procedure using writing test scores and Scholastic Aptitude Test scores to estimate more accurately the number of students needing placement in basic composition courses and now holds classroom assignments for them.

To learn more about the IUB campus assessment plan, please contact Deborah Olsen, Office of Academic Affairs and Dean of the Faculties, Bryan Hall 109, Indiana University, Bloomington, IN 47401. Tel.: (812) 855-1201.

References

Campus Assessment Plan. Bloomington: Office of Academic Affairs and Dean of the Faculties, Indiana University, 1995.

Farr, R., and Greene, B. *Indiana University Student Performance Measure*. Bloomington: College of Arts and Sciences, Indiana University, 1994.

Wolf, B. *Assessment Handbook*. Bloomington: Office of Academic Affairs and Dean of the Faculties, Indiana University, 1993.

Assessment Update
November-December 1995
Volume 7, Number 6

Assessment of Student Learning and Development at Winthrop University: An Update

Guest Columnists: Joseph Prus, professor of psychology and director of the Office of Assessment at Winthrop University, Rock Hill, South Carolina; and Margaret Tebo-Messina, associate professor of English at Winthrop University and director of the South Carolina Higher Education Assessment Network.

Winthrop University created its Office of Assessment and embarked on a comprehensive program to assess student learning and development in 1988. The foundations for the program included a set of principles established by the faculty to guide assessment; strong administrative support for a faculty-driven, multi-method approach; faculty and student advisory boards to represent various university constituencies and units; and keen faculty commitment to teaching and student learning. Under the administrative auspices of the vice president for academic affairs, the Office of Assessment was given the responsibility of initiating and coordinating universitywide activities such as general education assessment; student, alumni, and faculty surveys and interviews; and assessment of student personal growth and development. This support unit, staffed by teaching faculty on part-time release, was also created as a resource for department assessment efforts, such as those related to assessment of majors.

Although much has changed at Winthrop since our overall assessment activities were last reported in the *Campus Profile* column of *Assessment Update* in 1992, the same basic principles and structures for assessment remain in place. This report summarizes some of the key changes and developments that have occurred in our assessment program over the past few years.

Two assessment advisory boards, one representing students and the other representing faculty and senior staff, were created during the very early stages of program planning and development.

Assessment Advisory Boards

Two assessment advisory boards, one representing students and the other representing faculty and senior staff, were created during the very early

Assessment Update
November-December 1996
Volume 8, Number 6

stages of program planning and development. The faculty and staff board has been instrumental in representing the interests and needs of various academic units and in providing both philosophical and pragmatic direction. It reviews Winthrop's annual *Plan for Institutional Effectiveness* and *Institutional Effectiveness Report*, both of which are submitted annually to the South Carolina Commission on Higher Education. The board has also conducted periodic reviews of the Office of Assessment and helped oversee a major project in which academic units transformed their student learning goals and objectives and assessment plans into a common format and level of specificity. In 1995-1996, the board's responsibilities were broadened to include both assessment and institutional research, which are coordinated by separate academic support departments that collaborate closely.

The student advisory board was very effective early in the development of our program: it provided input from the student perspective and suggested means by which to reach, educate, and motivate students on assessment. One visual arts major on the board designed an assessment brochure that was subsequently printed and distributed to all students asked to participate in general education assessment activities. Other board members helped to write articles on assessment for the student newspaper. Once the newness of assessment at Winthrop faded, however, it became increasingly difficult to engage student board members through advisory activities. Thus, in 1993-1994, a decision was made to transform the student board into an active assessment group and the Students at Winthrop Assessment Team (SWAT) was born.

A decision was made to transform the student board into an active assessment group and the Students at Winthrop Assessment Team (SWAT) was born.

SWAT includes 8 to 12 students and a faculty and graduate student adviser. The students are recruited from disciplines for which assessment methods such as interviewing are particularly useful skills (for example, psychology, social work, marketing, mass communication). They receive 1 to 3 academic credit hours (depending on level of participation) for independent research. Those who agree to continue for a second year as leaders receive a small stipend from the Office of Assessment. The team focuses on issues of particular interest to students and on assessment projects that might be more effectively conducted by students (for example, interviewing fellow students in campus residences and in the halls of classroom buildings).

The team's completed projects have addressed student attitudes and campus experiences regarding human and cultural diversity, student behaviors and practices that lead to academic success, and student views on the characteristics of effective courses and instructors. The results of this last project served as the basis for a very successful student-led workshop for faculty. This past year SWAT collected survey data from students, parents, faculty, alumni, and community leaders and policymakers on what personal characteristics the college experience typically influences and which can or should be addressed in the goals of colleges and universities. This study, fashioned loosely after a study on critical thinking goals conducted by the National Center on Postsecondary Teaching, Learning, and Assessment, should provide useful information for examining the personal growth and development goals addressed in Winthrop's mission statement.

Assessment Update
November-December 1996
Volume 8, Number 6

General Education Assessment

Winthrop has continued to develop a multi-method interdisciplinary approach to assessing general education that focuses on first-year students and seniors as well as alumni. In addition, during the past two years the General Education Committee, in consultation with the Office of Assessment, began an analysis of the entire curriculum and extracurricular university experience to determine when, where, and to what extent students had opportunities to acquire general education skills and abilities.

Assessment measures include surveying seniors and alumni and evaluating first-year and senior skills and knowledge with a national standardized test and a collection of writing placement samples and writing-for-a-grade in Class Portfolios. Winthrop's ongoing efforts to develop institution-specific measures include several interdisciplinary faculty groups: the Oral Communication Committee, which was created in 1994; the longstanding Writing Assessment Committee, which oversees the collection and evaluation of writing samples and has surveyed faculty to determine the types of writing they require; and the six-year-old Learning Research Project, which is composed of classroom research faculty teams.

In 1994, Winthrop's Learning Research Project was honored as an exemplary assessment program by the South Carolina Commission on Higher Education and the South Carolina Higher Education Assessment Network. The project recruits and trains faculty in the use of classroom assessment techniques and uses an interdisciplinary team approach to develop program-level assessment measures. To date, approximately 30% of Winthrop faculty have taken part in the program, and they have assessed the learning of more than 4,000 students at every grade level and in every school and college of the university. In addition, the teams, which are focused on a general education goal of their choosing, have worked to create program-level assessment measures. For example, the Aesthetics Team has spent three years on a metaphor project, creating teaching and assessment materials, including an interdisciplinary video, with which to assess students' aesthetic appreciation.

Winthrop's Learning Research Project was honored as an exemplary assessment program.

Surveys and Interviews

During the early stages of assessment program development, Winthrop created the Follow-Along and Career-Tracking System (FACTS). The system included an entering student survey, withdrawing student survey, senior survey, one-year postgraduation alumni telephone interview, and three-year postgraduation alumni mail survey. Instruments included items pertaining to the quality of academic services and preparation, participation in extracurricular and postgraduate activities, and career and education goals and experiences.

Due to financial constraints experienced by the Office of Assessment in 1993 (one grant ended and a state budget shortfall occurred), the entering student survey was eliminated and the responsibility for the withdrawing student survey was moved to another department responsible for the university's retention plan. Additionally, it was necessary to complete the alumni telephone interviews and mail survey in alternate years rather than conducting both each year.

Responses of students and alumni to the various instruments included in Winthrop FACTS have been consistently positive, as have faculty, staff, and administrator responses to the results. Response rates for the senior survey, which was incorporated into an exit conference with each graduating senior conducted by the office of Records and Registration, have been over 90% consistently. Response rates for our computer-assisted alumni interviews have been about 50% of each year's graduating class of about 700 to 800 students, or about 75% of those for whom we have valid, working telephone numbers one year after graduation.

An executive summary of results for each survey or interview is mailed to all campus offices and departments. Customized reports with results for each major, including information on alumni (for example, employers, job titles, and graduate schools attended), are also mailed to each academic department. Executive summaries, which now include five-year retrospectives, and department reports are consistently evaluated quite positively by faculty, staff, and administrators and have prompted numerous changes. The Department of Career Services, for example, has completely revamped its services in response to student and alumni feedback.

Winthrop's faculty-centered, process approach to assessment has yielded many positive results.

Conclusion

Sustaining the growth of these multiple projects has proved challenging for a variety of reasons. Resources for higher education have shrunk and accountability demands expanded, increasing faculty frustration and decreasing faculty ownership of assessment. The university's general education governance structure has remained static and many problems inherent to program-level assessment have gone unresolved. For example, general education sample sizes are often too small to be statistically significant: too few seniors participate in the standardized testing and too few papers are contributed to the first-year and senior Class Portfolios.

Despite these problems, Winthrop's faculty-centered, process approach to assessment has yielded many positive results. Data gathered from the variety of methods currently in place have been widely disseminated and critiqued in terms of their validity and potential implications for curriculum and instruction. In the area of majors, for example, faculty have become more concerned with the quality of assessment methods and how their results lead to program improvement. Both the data and the process have helped create a campus culture that is generally accepting of assessment.

Assessment Update
November-December 1996
Volume 8, Number 6

PETER J. GRAY

Truman State University

When Truman State University was first featured in the *Campus Profiles* column in 1991, it was called Northeast Missouri State University. At that time the university had been experimenting with assessment for over 15 years. Over this period it had been transformed from a regional state teachers' college to a regional comprehensive university to a public liberal arts and sciences university. Part of that transformation was the refinement and adaptation of assessment techniques to be compatible with the changing mission of the new university. What began in the 1972-1973 academic year as a focus on nationally standardized examinations in the major had grown by 1975 to include multiple assessment measures such as surveys of students' satisfaction with academic and personal achievement and a program of general education value-added testing.

An especially important and unique feature of the Truman assessment program is the participation of every student. This ensures that assessment results are useful to individual students in monitoring their areas of strength and weakness and to the university for continuous improvement of its programs. By using assessment results in the advising process, the university has ensured their relevance to students and thus secured continued student participation.

In looking back over this history of experience, we can clearly see that assessment information by itself was not sufficient for transforming an institution. It was necessary to create a "campus culture that not only embraces assessment conceptually but actively supports organizational self-evaluation and the desire for improvement" (Magruder, McManis, and Young, 1997, p. 7).

Four factors supported this change in the campus culture. First, creation of the culture depended on the continuous commitment of the university's leaders, who clearly articulated their vision to the faculty. Over time a sense of trust was built that assessment was intended to monitor the institution's efforts to achieve student learning goals rather than to evaluate individual

An especially important and unique feature of the Truman assessment program is the participation of every student.

Assessment Update
September-October 1997
Volume 9, Number 5

faculty performance. Trust was fostered through one-on-one conversations with faculty members about specific students; by not making comparative data on majors or faculty members public; and by having faculty, the president, and the vice president for academic affairs use assessment data at workshops, planning retreats, and governing board meetings to analyze the strengths and weaknesses of the campus, to set goals for the future, and to consider the establishment of new programs. These activities also cultivated faculty ownership of the assessment process and the resulting information.

The second factor in establishing the campus culture of assessment was timing. When the accountability movement began in the early 1980s, the university was ready with objective measures of the quality of its education programs. By having adopted assessment as an institutional value, the university was not subjected to pressures from external mandates. And the university's early adopter role put it in a position of local and national leadership. All of these circumstances and the call for assessment issued in such reports as *A Nation at Risk* reinforced the university's commitment to focusing on student learning outcomes.

Because faculty have been directly involved in the development and implementation of the assessment program at Truman, they are very supportive.

The commitment of campus leaders to assessment and the reinforcement provided by external events provided the foundation for the third factor in establishing the culture of assessment: the integration of assessment into essential university processes. For example, an annual planning review defines more broadly the student learning process and the range of institutional performance measures. Assessment results have been integrated into the ongoing management and operation of the university.

Over the years three areas were identified as needing improvement. One area was the general education requirements for mathematics and science. A second area was student writing skills. The third was expectations related to course requirements in the majors that translated into increased student time on task. Emphasizing these types of internal improvements, rather than writing reports to comply with state mandates, has helped to make assessment relevant to faculty and students.

The fourth factor in the transformation of the campus culture was the direct involvement of faculty in the development and implementation of assessment practices. As noted by Magruder, McManis, and Young (1997, p. 4), faculty began in the mid 1970s to develop their own instruments for assessing student learning: "Over the following two decades, faculty have conducted a continual search for the best mixture of assessments—national and local, qualitative and objective—that would provide the university with multiple, appropriate measures of students' knowledge, skills and attitudes." Faculty also used assessment data for program review and they participated as readers for liberal arts and sciences portfolios and for the universitywide sophomore writing assessment.

Assessment Update
September-October 1997
Volume 9, Number 5

Because faculty have been directly involved in the development and implementation of the assessment program at Truman, they are very supportive, especially compared to faculty at other, similar institutions. In addition, their support has resulted in many changes in the basic nature of the student learning process at the university. There has been an increase in faculty-student interaction, in expectations for student achievement, and in students' active involvement in their learning, especially in small introductory classes. Students report spending more time studying and more students go on to graduate and professional schools. Learning communities have been established where faculty and students meet outside the classroom, and residential colleges have been created to provide living-learning environments.

Assessment results have helped faculty to integrate within the majors instructional activities focused on basic writing, speaking, and higher-level thinking skills. Similarly, the liberal studies program has been updated to promote cognitive skills development. In addition, a program of undergraduate student research was established, along with a first-year student transition program.

It is clear from the experience at Truman that the culture of assessment needs constant nurturing and care. As students, faculty, and administrators come and go, there is a need to acclimate people continually to the assessment culture. This process, too, must become part of the campus culture. Only in this way will it become self-sustaining.

To learn more about assessment at Truman, contact Candace Young, professor of political science and president of Truman's faculty senate, Truman State University, Kirksville, MO 63501-4221. Tel.: (816) 785-4650, E-mail <cyoung@truman.edu>.

As students, faculty, and administrators come and go, there is a need to acclimate people continually to the assessment culture.

Reference

Magruder, J. M., McManis, A., and Young, C. "The Right Idea at the Right Time: Development of a Transformational Assessment Culture." In P. J. Gray and T. W. Banta (eds.), *The Campus-Level Impact of Assessment*. New Directions for Higher Education, no. 100. San Francisco: Jossey-Bass, 1997.

Broadening the Scope of Assessment at Ball State University

In the early 1990s the focus of assessment at Ball State shifted away from a value-added emphasis to a more comprehensive approach to assessing the educational environment and its impact on students.

Assessment Update
January-February 1998
Volume 10, Number 1

The first Campus Profile of Ball State University appeared in this column in 1991. At that time Ball State had already experienced several major changes in relation to assessment. The first was in 1986, when Ball State's mission statement was amended to call for constant and vigorous self-assessment. Concurrent with the change in the mission statement was an effort to revise the university's general studies program. In addition, an office of academic assessment was established in 1987. This office works in parallel with the university's office of institutional research and reports to the associate provost.

Together these events provided the foundation for the work described in the earlier profile. In this profile I briefly summarize those early activities and describe the progress made at Ball State since 1991 (based on Palomba, 1997).

Initially, the office of academic assessment focused on the assessment of general studies and the collection of institutional data. Experience was gained with several national comprehensive examinations for general studies, including the American College Testing's College Outcome Measures Program exam and the Academic Profile. Through consultation with several departments, assessment was introduced in the disciplines. As a result, a number of departments used the Student Goals Exploration survey from the National Center for Research to Improve Postsecondary Teaching and Learning, which allowed teachers to determine learning goals of students in general education classrooms. In addition, six departments began to use Major Field Achievement Tests from the Educational Testing Service.

Based on these experiences, in the early 1990s the focus of assessment at Ball State shifted away from a value-added emphasis to a more comprehensive approach to assessing the educational environment and its impact on students.

In fall 1991, the provost and vice-president for academic affairs stated the following overall goals of Ball State's academic assessment program: to evaluate academic programs and to enhance student learning. Assessment was described as providing opportunities to demonstrate the university's strengths as well as to identify areas for growth and improvement. All colleges, schools, and departments were expected to participate in assessment activities and to use the results to stimulate program improvement. Faculty, administrators, and students were expected to work together to accomplish assessment goals. And the role to be played by alumni and community representatives in providing information and support for assessment activities was identified.

The provost articulated six major assessment objectives that have guided the activities at Ball State since 1991. These include determining the knowledge and attitudes of students at three points: when they enter the university, when they complete the general studies program, and when they finish their major. Other objectives include determining the factors that contribute to program completion, determining students' satisfaction with their educational experiences, and assessing students' success in employment and further education.

Six major assessment objectives have guided the activities at Ball State since 1991.

Along with a supplementary document describing assessment in each of the colleges and schools at the university, a modified version of the plan that resulted from the provost's statement was submitted to the North Central Association of Colleges and Schools during the fall 1993 reaccreditation visit.

Each of the provost's six objectives is addressed through one or more assessment activities. Standardized testing and surveys are used to determine the initial knowledge and attitudes of students and how these change over the course of study. Surveys are also used to determine the satisfaction and success of students at specific points during their academic experiences. The assessment activities that have been under way since 1993 fall into three broad categories: university-wide, general education, and discipline-specific.

University-wide assessment activities are used to address overall issues of learning that are important to the university, to help the units of the university see common goals, and to create university norms. The results of these activities are analyzed in a way that provides discipline-specific information. The university-wide assessment activities concentrate on learning objectives that cut across discipline lines and address broad learning objectives, such as critical thinking, clear communication, and the ability to work in groups. They also are used to shed light on the important areas of wellness, computer competence, and experiential education.

Surveys are administered to students at various times while they are at Ball State and after they leave. Information is collected from students when they enter the university, at midcollege, upon graduation, and two years after graduation. A common core of questions included in all surveys, and each survey also focuses on issues relevant to the student at that point in his or her school career. For example, each fall entering students complete the Making Achievement Possible (MAP) survey, developed at Ball State. This survey asks students to provide a self-assessment of their knowledge and skills; to indicate their willingness to seek help; and to describe their college goals and life goals, their initial adjustment to college, their plans for participating in college activities, and the ways they are likely to use their time. The data collected through MAP have been used to conduct retention studies that have increased understanding of the characteristics and behaviors contributing to program completion.

The College Basic Academic Subjects Examination has been used since 1989 to assess general education because it provides a variety of scores for English, mathematics, science, and social studies as well as provides scores for three levels of reasoning. This exam is used to study the knowledge of entering freshmen and their growth through the general studies program.

Another general education assessment has been in place since 1987, when the university implemented a required Writing Competence Examination for juniors that focuses on whether the student can take a position and defend it. This examination is conducted in a special two-hour testing session; students must write an essay in response to a prompt given at the session. Essays are graded on a three-point holistic scale.

Several elements have facilitated the assessment work at Ball State over the last ten years . . . One of the most important has been the consistent leadership for assessment at the top.

Recently a five-year cycle of general studies course assessment that began in fall 1991 was completed. Based on an established calendar and on reporting guidelines for assessment, each department assessed its own general studies courses and prepared a report for the general studies subcommittee of the university senate. Departments with at least one general studies course were encouraged to develop their own assessment tools to rate the university's programmatic goals for general studies and to provide both cognitive and attitudinal information from students to demonstrate that each course was meeting the general studies goals.

Another aspect of the general studies course assessment was the review of course syllabi by members of the General Studies Subcommittee. As a result of the review, the subcommittee has requested that syllabi (1) demonstrate that general studies courses are part of an overall program rather than a series of isolated courses, (2) emphasize programmatic goals, and (3) contribute to the overall coherence of the program.

The seven colleges at Ball State have adopted their own discipline-specific approaches to assessment. As a result, a wide variety of activities are in place, including portfolios, standardized tests, surveys, focus groups, classroom assessment, and other methods. The activities chosen by the colleges and departments and carried out by these units, often with support from the office of academic assessment.

Several elements have facilitated the assessment work at Ball State over the last ten years and especially since its last Campus Profile appeared in 1991. One of the most important has been the consistent leadership for assessment at the top, within colleges, and within departments and special programs. Another helpful element has been the point of view that university-wide

Assessment Update
January-February 1998
Volume 10, Number 1

assessment activities should be useful to the disciplines. A third element has been the expectation that assessment results should be reported and shared in varied and meaningful ways. A final element has been the recognition that if faculty are expected to participate in assessment activities, they should be provided appropriate support in terms of faculty development and reward.

To learn more about assessment at Ball State University, contact Catherine A. Palomba, Director of the Offices of Institutional Research and Academic Assessment, Ball State University, Muncie, Indiana, 47306-0220 (503-285-3716).

Reference

Palomba, C. A. "Assessment at Ball State University." In P. J. Gray and T. W. Banta (eds.), *The Campus-Level Impact of Assessment: Progress, Problems, and Possibilities*. New Directions in Higher Education, no. 100. San Francisco: Jossey-Bass, 1997.

If faculty are expected to participate in assessment activities, they should be provided appropriate support in terms of faculty development and reward.

Student Outcomes Assessment at Ohio University

Ohio University has been using student assessment information in program review, curriculum planning, and academic program planning.

Although the state of Ohio has never mandated assessment in any form, since 1981 Ohio University has been using student assessment information in program review, curriculum planning, and academic program planning. As described in Williford (1997), academic colleges, planning groups, individual departments, and trustees rely on outcomes information for both short-term and long-term planning and decision making.

Ohio University's progress in enhancing quality has been documented through its Institutional Impact Project (Williford and Moden, 1993), which includes a variety of quality indicators related to students' experiences. Changes can be measured by tracking these indicators over time. By being related to the university's goals and planning processes, the evaluative information embodied in the quality indicators has proved useful in guiding university improvement.

Williford (1997) describes two phases of Ohio University's experience with assessment. First, university-wide assessment information was provided to faculty and staff. In the second phase, individual academic units were supported in their need for assessment information.

The goal of the first phase of assessment was to provide reliable data to evaluate the performance of individuals and programs in achieving their goals. An essential element of the first phase was the Institutional Impact Project that was designed to address such university-wide goals as "providing our students with the knowledge and skills which are the essence of a solid liberal education" and "encouraging the development of an environment on the residential campus that reflects a vital commitment to learning and provides a community life for students" (Ohio University, 1977, pp. 18, 38).

The original Institutional Impact Project had five components. The first was the ACT College Outcome Measures Program Objective Test, which was used to assess Ohio University's general education program. Second, student tracking, retention, and graduation rate data were used to provide information

Assessment Update
March-April 1998
Volume 10, Number 2

BEST COPY AVAILABLE

useful to the university's retention programs. Third, student surveys were used to assess how students perceive the way they are treated by faculty and staff and to identify the activities in which students are involved. Fourth, a freshman marketing survey for admitted first-year students asked why they apply for admission to and enroll at Ohio University. Fifth, two follow-up surveys of graduates (one year and five years after graduation) yielded student outcomes information that was of interest to central administrators and to those in individual colleges and departments.

The information gathered through the components of the Institutional Impact Project has been used in many ways. For example, assessment information tied closely to the university's strategic and long-range plans guides the improvement of institutional quality. Assessment information also is used in the university's ongoing program planning. The university's curricular review processes incorporate information about student outcomes in making judgments about academic programs. Academic program review assessment data are used when departments are reviewed every seven years by the university curriculum council. Assessment information has been used to secure external funding for general education and for program-specific awards (such as Program Excellence, a part of the Ohio Board of Regents' Selective Excellence program). Having more than ten years of assessment information also has enabled Ohio University to respond positively to statewide reviews of efficiency and program quality.

Assessment information has been used to secure external funding for general education and for program-specific awards.

Although the original Institutional Impact Project was conceptualized with faculty and staff involvement, its operation was carried out through the efforts of Institutional Research staff with funds provided by the provost's office.

The second phase of assessment at Ohio University began with a 1993 review by the North Central Association of Colleges and Schools (NCA). During this review it became apparent that faculty were not involved in student assessment. As a result, the NCA required Ohio University to redraft its assessment plan to focus on department-based assessment, which provided a needed incentive to persuade faculty to be more involved.

In 1993-94, the provost asked Ohio University's academic deans to work with their faculty and the Office of Institutional Research to draft department-based assessment plans intended to improve teaching, learning, and student services. Consequently, departments developed learning objectives for their students; devised methods to assess the objectives using existing assessment methods and new methods; identified faculty or groups of faculty responsible for assessing students; set an implementation time line; and articulated possible uses for the assessment information.

Departments were given one ten-week academic quarter to develop their assessment plans with support and training provided by the Institutional Research staff. Although some colleges and departments viewed developing an assessment plan as a positive opportunity for program improvement, some resentment about developing an assessment plan for the NCA surfaced.

However, once faculty began to realize the institutional commitment to and the rewards of assessment, they became more involved and committed.

With the development of the new assessment plan in 1995, student assessment was organized at three levels: institution-wide assessment, institution-wide support of department-based assessment, and department-based assessment. In March 1995, the president of Ohio University announced that the university would renew its commitment to assessment and would implement department-based plans beginning in the 1995-96 academic year. A communication strategy was designed to hold faculty accountable for departmental assessment. The strategy involved having each department prepare a brief, four- to eight-page report on assessment activities and results. These reports were shared among the colleges for comment, review, discussion, and implementation. From the department, a summary report went to the dean of the college for review and comments. Then the summary report was either returned for revisions or forwarded to the provost.

Once faculty began to realize the institutional commitment to and the rewards of assessment, they became more involved and committed.

At the end of the 1996 spring quarter, departments submitted their first assessment reports: brief summaries of how their original plans had been implemented and program improvements made or planned. Based on the first set of reports, the following list of good practices in assessment was created (Williford, 1997, p. 54): a clear statement of department-specific goals that matches reported assessment activities; faculty involvement in curricular assessment and improvement; use of multiple measures for assessment data; use of information already gathered and distributed by the Office of Institutional Research; integration of departmental, college, and university missions; focus on student outcomes with emphasis on both benchmarks and value-added measures; improvements based on results or a plan of how results will be used for improvement; assessment activities or other elements of the model that are generalizable to other units; and a continuum of assessment activities from students' first year to their after-graduation work experiences.

The curricular review process also was redesigned in 1995-96; instead of requiring departments to provide documentation on the strength of their curricular and faculty resources, the new review process asks departments to provide evidence that they have improved the quality of their programs—teaching and learning—using student assessment information.

The Institutional Impact and Assessment Plan created in 1994-95 is a flexible strategy in that it recognizes that assessment objectives, activities, and uses will change over time. Since the primary purpose of assessment at Ohio University is to assist the institution's units in enhancing program quality, a current goal is to enable faculty to become involved in assessment so that they will know more about and become more involved with their programs and their students. However, there are two other important goals for the future. One is for students to become more involved in assessment so that they will know more about themselves and become more involved in their

Assessment Update
March-April 1998
Volume 10, Number 2

education. The second is to address issues of quality in creative and scholarly activity and in service to the community.

At Ohio University assessment is seen as providing a framework for improving the quality of the full range of faculty and student activities related to teaching, learning, research, and service.

For more information about assessment at Ohio University, see the Williford chapter in Gray and Banta (1997) or contact: A. Michael Williford, Director of Institutional Research, Ohio University, Athens, OH 45701, (614) 593-1059.

References

Gray, P. J., and Banta, T. W. (eds.). *The Campus-Level Impact of Assessment: Progress, Problems, and Possibilities*. New Directions for Higher Education, no. 100. San Francisco: Jossey-Bass, 1997.

Ohio University. *Ohio University Educational Plan: 1977-1987*. Athens: Office of the President, Ohio University, 1977.

Williford, A. M. "Ohio University's Multidimensional Impact and Assessment Plan." In P. J. Gray and T. W. Banta (eds.), *The Campus-Level Impact of Assessment: Progress, Problems, and Possibilities*. New Directions for Higher Education, no. 100. San Francisco: Jossey-Bass, 1997.

Williford, A. M., and Moden, G. O. "Assessment and Quality Enhancement at Ohio University." In T. W. Banta (ed.), *Are We Making a Difference? Outcomes of a Decade of Assessment in Higher Education*. San Francisco: Jossey-Bass, 1993.

Assessment is seen as providing a framework for improving the quality of the full range of faculty and student activities related to teaching, learning, research, and service.

Institutional Effectiveness at Midlands Technical College

The progress on goals and initiatives is reported annually in an institutional effectiveness report card.

Midlands Technical College (MTC) was first profiled in spring 1993 (Gray, 1993—page 330). At that time the college, which serves the Columbia, South Carolina, metropolitan area with three campuses, had seen a 61 percent enrollment growth since 1978. Regular enrollment was eleven thousand and continuing education enrollment stood at fifteen thousand. Nearly one-quarter of the 1990 area high school graduates who went on to college enrolled at MTC.

In 1986, the faculty, staff, and governing board of MTC initiated a strategic planning process that provided the foundation for a comprehensive institutional effectiveness initiative that began in 1988. The strategic planning process was used to identify the major direction and priority initiatives of the college, beginning with clarification of the college's mission, its role and scope, and its values.

Based on these considerations, institutional goals were created for the period from 1992 to 1997, as described in *A Vision for Excellence* (Midlands Technical College, 1993a). The eight original goals set in 1992 were modified into nine goals after a comprehensive evaluation of the college's mission, strengths, and opportunities, and long-range direction was undertaken in the 1994-95 academic year (Midlands Technical College, 1995b).

Each of the nine college goals advances several priority initiatives. As part of a three-year strategic planning cycle, the progress on these goals and initiatives is reported annually in an institutional effectiveness report card.

The 1996-97 report card (Midlands Technical College, 1997c) describes progress achieved on each of sixty-six priority initiatives. For example, the first priority initiative under goal one ("to revise, enhance, and develop the curriculum") was to "develop or reconfigure curricula as appropriate to expand educational options for students" (p. 2). Eleven different sets of activities were described that detailed the progress achieved on this priority

Assessment Update
September-October 1998
Volume 10, Number 5

BEST COPY AVAILABLE

initiative. These activities included ten new non-institutional, community-based clinical experiences for health sciences students and a 102-hour auctioneer licensing credential program in the Continuing Education Division (p. 2).

College goal four in the 1996-97 report card was to "support the college's commitment to access, equality, and diversity" (p. 7). Toward that end a priority initiative was established to "implement a professional development program for faculty and staff that provides the necessary support for working with students who have special learning needs or disabilities" (p. 9). Progress related to this initiative included the design of a "Putting People First" disabilities awareness seminar piloted by directors and program managers of student development services. The seminar received excellent evaluations and it was implemented for faculty and staff during 1997-98 (p. 9).

A college goal that was modified for the most recent period was "to seek additional resources to support the college's mission" (p. 13). The first priority initiative under this goal was to "increase private foundation and alumni support and continue to pursue public and private grant funds." Under the progress achieved in relation to this initiative were sixteen examples of funding that ranged from $3,800 donated to a scholarship fund to $1,975,410 in federal funds secured to support job training (pp. 13-14).

For each situation in which there is a discrepancy between the standard and the results, an action plan is proposed.

In addition to the college goals and priority initiatives reviewed each year of a three-year planning cycle, there is a companion process related to critical success factors. Each critical success factor is described by indicators of effectiveness and measured by a set of standards. There are six critical success factors: (1) accessible, comprehensive programs of high quality; (2) student satisfaction and retention; (3) post-education satisfaction and success; (4) economic development and community involvement; (5) sound, effective resource management; and (6) dynamic organizational involvement and development.

The 1996-97 report card also describes the standards and results related to each indicator of effectiveness. For example, although the access and equity results show that the data on enrollment by race meet the standard of being "within 5 percent parity with the composition of the community," the data on enrollment by gender do not (p. 19). The interpretation offered is that "enrollment patterns based on gender are not within the established standard. They are reflective of national enrollment patterns" (p. 19). For each situation in which there is a discrepancy between the standard and the results, an action plan is proposed. For example, to address the discrepancy between the gender equity standard and the current results, it is proposed that the college "market programs to industry to capture males already in the work force" (p. 19). The data collected through the critical success factors process provide valuable information for those who are making decisions about the future of the college.

The most recent operational plan lays out the action strategies related to each college goal and priority initiative for 1997-98 based on the results described in the 1996-97 report card (Midlands Technical College, 1997b). The plan demonstrates the college's ongoing commitment to accountability.

Strategic planning has been incorporated into the self-study process at MTC. As a result, (1) a broad spectrum of the college's population will provide direct input into developing the next strategic plan, (2) the most current data available will be used to evaluate the college's present status, and (3) the self-study process will be used to enhance the effectiveness of the institution (D. A. Kitchings, personal correspondence, April 1998).

An essential element of the self-study process was a SWOT (strengths, weaknesses, opportunities, and threats) analysis. Using the following questions, the self-study subcommittees identified specific factors that will influence the future growth of the college (Midlands Technical College, n.c.b.):

How do we accommodate the demands of the new learner who wants to learn anytime and anywhere?

What are the anticipated changes in the environment that will significantly affect the organization and development of higher education?

What do we need to do today to position ourselves to successfully compete in the future?

How do we prepare ourselves for the transformation to the knowledge age?

What are the existing barriers to our transformation in the future? What are some strategies to overcome those barriers?

What future organizational structure will be required to facilitate growth?

What new skills should our workforce (faculty and staff) possess in the future?

What technical infrastructure will be needed to facilitate growth?

What are alternatives to current learning models and what new academic cultures will we need to develop for the future?

How do we accommodate the demands of the new learner who wants to learn anytime and anywhere?

What new strategic alliances and partnerships do we need to develop in the future?

The self-study will go beyond the traditional compliance model and will have a special focus on "envisioning the future" (D. A. Kitchings, personal correspondence, April 1998), which will direct the process of planning and prioritization into the next decade. Three essential questions will be used to guide this visioning process: (1) What have we learned over the last ten years that we should continue to do into the next decade? (2) What changes will make us better? and (3) What directions do we need to take in the future?

A call for transformational leadership has been issued that is intended to position MTC strategically for the twenty-first century (Midlands Technical College, n.d.a.). In this call are described the risks related to either resisting or embracing the changes in the environment outside the college. A case is made for change by describing the factors in the environment that have implications for the college. The call concludes that "the risks of maintaining the status quo are more profound than the risks of changing" and that "incremental changes are not sufficient. Transformational change that affects the very nature and character of how we do business is needed" (p. 4). Finally, eleven actions are listed that must occur for change to be successful.

To help other institutions take advantage of the lessons learned over the last ten years at MTC, a user's guide for managing an institution's effectiveness has been created by MTC and published by the American Association of Community Colleges (Midlands Technical College, 1997c). The guide provides "instructions and representative forms or procedures that can be used by community colleges to construct their own institutional effectiveness . . . program" (p. vii).

A call for transformational leadership has been issued that is intended to position MTC strategically for the twenty-first century.

The user's guide is based on the view that institutional "effectiveness is not a measurement process, it's a change process" (Lorenzo, 1990). In the foreword to the guide, MTC's president, James L. Hudgins, depicts institutional effectiveness as a strategy for institutional renewal. He goes on to describe the why, what, and how of institutional effectiveness, which are embodied in the guide. The guide provides many examples of useful procedures, strategies, and forms, from the beginning of the institutional effectiveness process (step one, strategic planning) to the end (step five, using data to improve effectiveness). There is even a set of diskettes to permit the guide's figures and tables to be used as templates.

The guide describes in detail two areas in which previous assessment efforts at MTC have been enhanced. The first area is the academic program review process, which has been made more rigorous, as is made clear by the figures and tables in the guide. The second area is the academic advising process, which has been made more rigorous, as is made clear by the figures and tables in the guide. The second area is the academic advising process, which was recently recognized as an exemplary assessment effort by the South Carolina Commission on Higher Education.

The guide, along with the various planning documents and report cards that have been developed at MTC over the last ten years, attests to a sustained commitment to assessment, continuous improvement, and institutional effectiveness. For more information about MTC, please contact Dorcas A. Kitchings, Director, Assessment, Research and Planning, Midlands Technical College, P.O. Box 2408, Columbia, South Carolina 29202 (803-738-1400).

References

Gray, P. J. "Campus Profiles Midlands Technical College," *Assessment Update*, Vol. 5, No. 2, pp. 11-13, March-April 1993.

Midlands Technical College. *A Call for Transformational Leadership*. Columbia, SC: Midlands Technical College, n.d.a.

Midlands Technical College. *Questions Related to SWOT Analysis and Future Positioning for the Self-Study Subcommittees to Consider*. Columbia, SC: Midlands Technical College, n.d.b.

Midlands Technical College. *A Vision for Excellence*. Columbia, SC: Midlands Technical College, 1995a.

Midlands Technical College. *Strategic Planning Outcomes 1992-1995*. Columbia, SC: Midlands Technical College, 1995b.

Midlands Technical College. *Managing Your Institution's Effectiveness: A User's Guide*. Washington, DC: American Association of Community Colleges and Midlands Technical College, 1997a.

Midlands Technical College. *Operational Plan 1997-1998*. Columbia, SC: Midlands Technical College, 1997b.

Midlands Technical College. *Report Card 1996-1997*. Columbia, SC: Midlands Technical College, 1997c.

Assessment Measures

Gary R. Pike

About the Author

Gary R. Pike is assistant vice chancellor for student affairs and director of student life studies at the University of Missouri-Columbia (MU). Prior to coming to MU in 1993, he was the associate director of the Center for Assessment Research and Development at the University of Tennessee, Knoxville. In addition to serving as the editor of the *Assessment Measures* column in *Assessment Update*, Pike has served as an expert consultant for the U. S. Department of Education and has worked with two- and four-year colleges and universities in 14 states to develop and improve assessment programs. In addition to his *Assessment Measures* columns, Dr. Pike has written more than 25 articles and book chapters on assessment and student learning. He has also received three best paper awards from the Association for Institutional Research and received the Outstanding Assessment Research Award from the American College Personnel Association, Commission IX in 1998. In 1997, he served as a representative from the University of Missouri on the American Productivity and Quality Center's benchmarking study on assessment. Pike received his Ph.D. in communication from the Ohio State University in 1985.

Contents

		Page
Ten Years of *Assessment Measures*		421
Topic	*Title*	
General Education	Overview of General Education Assessment Instruments	429
General Education	Criteria for Selecting an Assessment Instrument	432
General Education	Joint Task Force of the Interinstitutional Committee of Academic Officers and State Board for Community College Education	434
General Education	The College Outcome Measures Program Objective Test	436
General Education	Academic Profile II	439
General Education	College BASE	442
General Education	Using Self-Report Measures as Proxies for Test Scores	445
General Education	Measuring Growth in Student Learning and Development	448
General Education	Freshman-to-Senior Gains at the University of Tennessee, Knoxville	452
Critical Thinking	Assessing the Critical Thinking Abilities of College Students	455
Critical Thinking	The Watson-Glaser Critical Thinking Appraisal	459
Critical Thinking	The California Critical Thinking Skills Test	462
Critical Thinking	The Reflective Judgment Interview	465
Critical Thinking	Critical Thinking Assessment at the University of Missouri, Columbia	468
Critical Thinking	Using Surveys to Measure Critical Thinking Outcomes	474
Surveys	College Student Experiences Questionnaire	479
Surveys	The Community College Student Experiences Questionnaire	481

Topic	*Title*	Page
Surveys	ACT Evaluation/Survey Service	484
Surveys/Test	ACT ASSET Program	486
Surveys	Student Reactions to College (SRC) Questionnaire	488
Surveys	The Noel-Levitz Student Satisfaction Inventory	491
Surveys	Student Goals Exploration	493
Surveys	The Consortium on Financing Higher Education's Alumni Survey	495
Surveys	The University of Missouri Freshman Survey	497
Surveys	University of Missouri's Survey of Admitted Students	500
Writing Assessment	Issues in Writing Assessment	503
Writing Assessment	Statewide Writing Assessment (K-12) in Missouri	506
Major Field	Assessment in the Major at UTK	510
Major Field	Project for Area Concentration Achievement Testing	513
Assessment Resources	Clearinghouse for Higher Education Assessment Instruments: Survey Results	516
Assessment Resources	Clearinghouse for Higher Education Assessment Instruments: Resources	519
Enrollment Management	Enrollment Management	521

GARY R. PIKE

Ten Years of *Assessment Measures*

I have to confess that I was somewhat taken aback when Trudy asked me to write a reflective essay to wrap around *Assessment Measures* columns written in the last ten years. First of all, it does not seem like it has been ten years since the first column was published in *Assessment Update*. Second, these columns have not had a consistent theme or themes over the years. I have tended to write columns that were related to interests of mine at the time, rather than having some grand plan for systematically surveying the waterfront of higher education assessment. In an effort to organize the *Assessment Measures* columns, I have provided a topical list in the Contents. Here I want to review briefly the topics covered in *Assessment Measures* and then share with you some of the thoughts that occurred to me as I was attempting to bring some coherence to this diverse set of essays and reviews.

The very first *Assessment Measures* column in 1989 reviewed nine objective tests that have been used to evaluate **general education outcomes**. Since that first column, the assessment of general education has been a frequent topic in *Assessment Measures*. Two other early *Assessment Measures* columns examined criteria that can be used in evaluating and selecting general education assessment instruments. The first of these, "Criteria for Selecting an Assessment Instrument," presents a set of criteria for evaluating general education measures that focus on the alignment of assessments with the goals of general education and the psychometric (i.e., measurement) properties of the instruments. The second column, a summary of the report of the Joint Task Force of the Interinstitutional Committee of Academic Offices and State Board for Community College Education, provides an example of how one state (Washington) employed a set of criteria to evaluate three well known tests of general education outcomes.

Three other *Assessment Measures* columns have provided brief profiles of three popular measures of general education assessment: *Academic Profile II*, *College Basic Academic Subjects Examination* (*College BASE*), and the *College Outcome Measures Project (COMP) Objective Test*. While these

***The assessment of general education has been a frequent topic in* Assessment Measures.**

BEST COPY AVAILABLE

profiles were not intended to be critical reviews of the tests, the essays do make use of the criteria for evaluating general education assessment instruments that I had outlined in an earlier column. Significantly, all three of these tests are widely used today.

The final three columns classified under the rubric of general education assessment focused on measurement issues. Two of these dealt with value-added, a concept that plays an important role in discussions of general education assessment. The column "Measuring Growth in Student Learning and Development," examined the strengths and limitations of three methods of representing the value added by general education, while the column on freshman-to-senior gains examines the value-added concept using data from the University of Tennessee, Knoxville. The third column dealing with measurement of general education outcomes addressed an issue that was central to discussions about a national assessment of college student learning: Can self-report measures provide useful information for evaluating the general outcomes of college—critical thinking, communicating, and problem solving? The answer to this question was an equivocal "yes."

The assessment of students' **critical thinking abilities** has been the subject of six columns in *Assessment Measures*. The first column in this series provided an overview of issues surrounding the assessment of critical thinking. As I reread that column today, two themes stand out. The first is Len Baird's model of college outcomes, which suggests that critical thinking abilities are content free and that these abilities are influenced by students' college experiences. The second theme, really a question, deals with the development of critical thinking abilities. Whether one views the development of critical thinking abilities as linear and additive or whether one views their development as progressing through stages, has profound implications for the types of assessment instruments and methods selected. It also admonishes assessment professionals that all instruments carry with them assumptions about the nature of what is being assessed and that failure to consider these assumptions may result in the selection of inappropriate measures.

*The assessment of students' **critical thinking abilities** has been the subject of six columns in* Assessment Measures.

The remaining articles dealing with critical thinking review a variety of types of assessment instruments ranging from objective tests—such as the *Watson-Glaser Critical Thinking Appraisal* and the *California Critical Thinking Skills Test*—to qualitative assessments based on expert judgment—the *Reflective Judgment Interview*. Phil Wood's column describes an attempt at the University of Missouri-Columbia to bridge the gap between objective measures and qualitative assessments, presenting an objectively scored assessment based on the reflective judgment stage model.

The final article in this set examines the use of self-report measures (e.g., surveys) to assess the development of students' critical thinking abilities. In addition to describing the types of information about students' critical thinking abilities that can be provided by various commercially available surveys, such as the ACT *College Outcomes Survey* and the *College Student Experiences Questionnaire*, the column also provides suggestions to assessment professionals interested in developing survey items that focus on the development of critical thinking.

The most frequently discussed topic in *Assessment Measures* columns to date has been the use of **surveys** in assessment research. Columns have profiled a variety of surveys, ranging from commercially available instruments (e.g., the *College Student Experiences Questionnaire*, the ACT *Evaluation/Survey Service*, and the *Noel-Levitz Student Satisfaction Survey*) to surveys developed at a single institution or by a group of institutions (e.g., the Consortium on Financing Higher Education alumni survey, the University of Missouri-Columbia freshman survey, and the *Survey of Admitted Students* also developed at the University of Missouri-Columbia). All of these columns reflect my belief that one survey is not inherently better than another. Each survey has its own advantages and disadvantages. Commercially available surveys offer the advantages of time to obtain and national comparison groups, while locally developed surveys offer the advantages of having questions tailored to specific institutional concerns and lower cost of administration.

The two *Assessment Measures* columns dealing with **writing assessment** came about as a result of my experiences with the statewide K-12 writing assessment in Missouri. The first column examines three issues in writing assessment: What sort of writing task is appropriate? What type of prompt should be used? What rubric (scoring scheme) should be used to evaluate the writing samples? The second column on writing assessment describes my experiences with the Missouri writing assessment program. I suggested three principles that should underlie writing assessment (and assessment in general). First, there is no substitute for rigor in assessment research. The second principle is that assessment frequently serves multiple purposes and involves multiple tradeoffs. The third is that assessment must be coupled with feedback to produce improvements in performance.

Many of the principles of assessment can and should be applied to research related to enrollment issues.

Although assessment of **learning in the disciplines** was the topic of the second column written for *Assessment Measures*, only one other column has dealt with this topic directly. This does not imply that assessment of learning in the disciplines is unimportant. Instead, the paucity of information about major field assessment in *Assessment Measures* reflects the breadth of assessments that are used in various disciplines, coupled with the fact that very little is written for publication about assessment in the major. Peter Ewell has described information about major field assessments as a "fugitive literature." What is needed, both in *Assessment Measures* and other publications, is to make this fugitive literature public.

Two of the three remaining *Assessment Measures* columns presented **resources** for assessment professionals. The first reported the results of a survey concerning types of assessment instruments used nationally and was conducted by the Clearinghouse for Higher Education Assessment Instruments. The second article described resources available from the Clearinghouse. The final topic covered in *Assessment Measures* dealt with issues related to enrollment management (e.g., recruitment, retention, and graduation). While this column was intended to be a springboard for discussions of instruments used in **enrollment management** research, it also suggests that many of the principles of assessment can and should be applied to research related to enrollment issues.

So what can be distilled from 32 columns dealing with diverse topics ranging from writing assessment in grades K-12 to complex statistical procedures involving multitrait-multimethod analysis? First, many of these columns either explicitly or implicitly suggest a set of criteria for evaluating the assessment instruments used on college campuses. These criteria focus on traditional measurement issues of reliability and validity, but they also focus on the alignment of test content with the goals of the institution. Alignment of test content with educational goals is important because experience suggests that the closer students' educational experiences correspond to the content covered by a measure, the greater the utility of assessment results. Stated plainly, if you want to know about the effectiveness of your education programs, you need to use measures that accurately represent the outcomes you expect from your programs.

If you want to know about the effectiveness of your education programs, you need to use measures that accurately represent the outcomes you expect from your programs.

In addition to content coverage, the psychometric properties of an instrument are critically important. There really is no substitute for good measurement. Most, if not all, assessment research contains a great deal of error. As I discussed in my columns on value-added, this error can create serious problems in assessment research when it comes to evaluating the effectiveness of education programs. Selecting dependable instruments that accurately measure what they purport to measure helps avoid compounding the measurement difficulties inherent in educational assessment, and it is essential if we are to make meaningful improvements in the quality of postsecondary education.

The final criterion for evaluating an assessment measure is its sensitivity to educational effects. Frankly, I fail to see how any assessment instrument can be useful if it does not accurately reflect the impact of students' educational experiences on learning and development. Unfortunately it is this criterion that poses the greatest hurdle for most assessment instruments. Experience and research suggest that student performance on achievement tests, whether commercially available or locally developed, is strongly influenced by general intellectual ability. Similarly, self-reports of students' college experiences seem to be subject to a constant error of the halo that obscures differences in students' learning and development across different outcome domains. The net result of both of these general factors is that identifying the outcomes of interventions designed to improve educational quality is extremely difficult.

While I still believe that all of the comments I have made about measurement goal alignment in the past are true, I also believe they do not go far enough. In recent months, staff in my office have been using a simple chart to relate the steps in the assessment process to steps in the research process and to the steps underlying the Plan-Do-Check-Act (PDCA) cycle of continuous quality improvement. Consistent with my admonitions, the process begins with the end in mind—the goals of the education program. However, it also includes questions related to what is being done, the appropriate research approach, and the types of evidence needed to evaluate the program. Only after these issues have been addressed do issues of method and analysis arise. In the final step of the assessment/research/improvement process, the chart emphasizes the importance of acting on assessment data.

Key Steps in Assessment, Research, and Quality Improvement Processes

The Steps Essential to Good Research	*The Steps Essential to Quality Improvement*
1. Choose a research question. Make sure the question is relevant to something important.	1. Decide what you want to accomplish. Make sure it is directly relevant to what the division wants to accomplish.
2. Choose an approach, i.e., explore what you can do to answer the question.	2. Choose an approach, i.e., explore what people are doing now, the programs and activities, to accomplish the goal.
3. Decide what evidence you need to answer the question.	3. Decide what evidence you need to decide how the programs and activities are working.
4. Choose and implement methods appropriate for gathering the evidence.	4. Choose methods appropriate for gathering the evidence.
5. Analyze the evidence and draw conclusions.	5. Analyze the evidence and decide how well things are working.
6. Report and communicate your conclusions.	6. Decide to continue, expand, revise or eliminate programs and activities based on your conclusions. Document what you learn and decide to do.

Simply looking at the alignment of program activities/curriculum with program goals is a basic assessment technique that can reap significant benefits without collecting a shred of data.

The problem with my earlier recommendations regarding the alignment of goals and test content is that they tend to encourage assessment professionals to jump from goals to methods without considering the intervening steps. Too often that involves running to select the test that best measures a program's goals. Stopping to consider questions about activities, approaches, and evidence can yield enormous benefits. For example, simply looking at the alignment of program activities/curriculum with program goals is a basic assessment technique that can reap significant benefits without collecting a shred of data. Connecting methods to program activities may also improve measurement. As I noted in my earlier comments on sensitivity to educational effects, tests and surveys tend to measure a broad range of outcomes at a fairly superficial level. If the program activities being assessed are focused initiatives, reliance on a test or survey may not provide sufficiently detailed information to assess program impact accurately. Stated directly, assessing changes in broad educational goals when improvement actions are focused initiatives is akin to hunting flies with a blunderbuss.

It has also been my experience during the past few months that thinking about what evidence is persuasive tends to shift the focus from traditional approaches, such as tests and surveys, to more qualitative approaches involving interviews of students and staff. This is particularly true when the focus of assessment is on the quality and effectiveness of specific programs or processes. The Missouri writing assessment program is both a good and bad example of selecting evidence that is appropriate for evaluating effectiveness. The strength of the writing assessment in Missouri is that both the nature of writing and the nature of writing instruction were taken into account in selecting a qualitative process-writing measure over an objective test or a one-shot sample of writing. The weakness of the Missouri writing assessment is that, while providing feedback to students and the state, it does not provide effective feedback that local school districts can use for improvement. Again, the culprit is a lack of specificity. Simply reporting holistic evaluations to school districts does not provide evidence about the effectiveness of specific aspects of the writing curriculum or of writing instruction. Only an analytic approach aligned with curriculum and instruction can provide that information.

Thinking about what evidence is persuasive tends to shift the focus from traditional approaches, such as tests and surveys, to more qualitative approaches involving interviews of students and staff.

In deciding what evidence is compelling, I find it helpful to visualize the final report I am writing and to honestly ask myself if the final report provides persuasive evidence of program effectiveness *and* suggests ways in which the program might be improved. If I were visualizing a report on writing assessment to local school districts, I could not honestly say that average holistic writing scores would provide persuasive evidence of a school district's writing program. Analytic scores, closely aligned with curriculum and instruction and disaggregated by student groups, would provide more persuasive evidence. Similarly, reporting that 63% of the students scored above the 50th percentile on a national examination of general education outcomes does not provide very persuasive evidence of a program's effectiveness. It certainly does not provide the type of evidence needed to guide improvement efforts.

I also find it helpful to think about a hierarchy of assessment evidence. Like Maslow's hierarchy of needs, the hierarchy of evidence progresses from a basic level of evidence about outcomes and the day-to-day functioning of the system to a more complex level focusing on relationships among pieces of evidence. The hierarchy next moves to consider the effects of system characteristics and ultimately focuses on the values underlying the system. In this model, the fact that 57% of the students scored above the national average on a test of mathematics achievement would represent the first level of evidence. The fact that a lower percentage of students were above the national average on mathematics achievement than on any other test subscore represents the second level in the hierarchy, as does the fact that motivation is strongly related to test performance and students' reported relatively high levels of motivation when taking the mathematics test.

Examining students' mathematics scores in light of how those students fulfilled their mathematics requirements (e.g., on campus, off campus, or exemption) would represent the third level in the hierarchy. It is also the level required for evidence to be persuasive, because only at this level is the linkage between program activities and educational outcomes explicitly made.

I have found that evidence at the values level is extremely rare, but I believe that it is essential if assessment is to produce meaningful institutional change. My belief in the importance of evidence linking outcomes to institutional values is based on the premise that the ultimate goal of assessment should be the creation of a learning organization, what Peter Ewell described as a *self-regarding institution*. The creation of learning organizations represents a change in institutional culture and values for many colleges and universities, and evidence showing how institutional culture and values help or hinder student learning and development can guide fundamental efforts to transform the institution. At the very least, evidence at the highest level of the hierarchy forces us to acknowledge the existence of various aspects of institutional culture and institutional values.

Effective assessment, like effective research and effective quality improvement efforts, requires a systematic process of inquiry and reflection.

In sum, effective assessment, like effective research and effective quality improvement efforts, requires a systematic process of inquiry and reflection. The process is systematic because all steps are always followed. The process begins with goals and programs and activities designed to meet those goals. Questions about approach and evidence are closely related and reflect a concern with information that can be used for improvement. Methods and analysis become secondary concerns in this model, while taking action on the basis of persuasive evidence is seen as the ultimate goal of the assessment process.

Being asked to review these *Assessment Measures* columns also taught me something about my own interests and biases. If you look at what have been written in *Assessment Measures*, it becomes apparent that I tend to focus on more global educational outcomes-student satisfaction, learning in general education, and the development of critical thinking abilities. At the same time, I tend to favor the use of self reports in assessment research. While global outcomes and self-report measures are important elements in a campus assessment program, they provide only part of the picture. Certainly, students' experiences in the major are at least as important as general learned abilities, and locally-developed measures are essential if we are to measure educational outcomes in certain disciplines. You would think that, having been at two institutions, each offering more than 100 different degree programs, I would have written more about assessment in the major. One reason I have been reluctant to write about locally-developed assessment at UTK and MU is that I have been concerned about giving too much attention to assessment at the institutions where I have worked.

I want to close this essay with a request for your help. Peter Ewell once characterized information on locally developed tests as a "fugitive" literature in outcomes assessment. At the same time, assessment in the major offers some of the greatest opportunities for meaningful and lasting improvements in the quality and effectiveness of higher education. Assessment in higher education has tended to evolve as a cottage industry in which each department and institution develops measures in isolation from the experiences of others. Assessment is a difficult and time-consuming process, and we should not try to reinvent the wheel every time we develop a new instrument. I am asking all of you to examine your own experiences, particularly your experiences with locally developed assessments in the major and to share your experiences with other by submitting essays to *Assessment Measures*.

W*e should not try to reinvent the wheel every time we develop a new instrument.*

GARY R. PIKE

Overview of General Education Assessment Instruments

◆

During the last two years there has been tremendous growth in the number of institutions interested in assessing the outcomes of general education and in the types of instruments available to assess those outcomes. A brief review of the literature on outcomes research reveals that prior to 1987 five tests were available for general education assessment: the ACT Assessment Program examinations, the College-Level Examination Program (CLEP) General Examinations, the College Outcome Measures Project (COMP) examination, the Graduate Record Examinations (GRE), and the Scholastic Aptitude Test (SAT). Since 1987, four new tests have been introduced: the Academic Profile, the College Basic Academic Subjects Examination (College BASE), the Collegiate Assessment of Academic Proficiency (CAAP), and the Education Assessment Series (EAS) examinations.

All but one of the tests in use prior to 1987 were designed for some purpose other than evaluating general education programs.

In this report I shall provide an overview of each of these tests that will include a brief description of each and an analysis of its strengths and weaknesses. Also included is a listing of the test publishers in case readers want to obtain additional information.

All but one of the tests in use prior to 1987 were designed for some purpose other than evaluating general education programs. For example, the ACT Assessment Program is a battery of college entrance and placement examinations. While research shows that students' scores on these tests predict performance during the first two years of college, there is no consistent evidence that the tests are sensitive to the effects of college coursework.

Likewise, the Scholastic Aptitude Test is designed for use as a college entrance and placement examination, and the quantitative and verbal subscales have been shown to predict performance in college. Publishers of the SAT describe these exams as measures of problem-solving ability; however, there is no evidence that they are particularly sensitive to the effects of general education coursework.

Assessment Update
Spring 1989
Volume 1, Number 1

BEST COPY AVAILABLE

The College-Level Examination Program (CLEP) General Examinations are designed to assess students' knowledge and skills in five content areas. While all five tests are highly reliable and have been linked to student performance in introductory courses in each content area, empirical data on the validity of the CLEP General Examinations as measures of program effectiveness currently are not available.

A test usually given at the end of the undergraduate years, the General Test of the Graduate Record Examinations (GRE), has also been used in assessing the effectiveness of the general education curriculum. The GRE is a nationally normed exam designed to measure skills that are gained over a long time period and are unrelated to a particular field of study. While this test is most often used to measure students' abilities for graduate study, recent studies conducted by the Differential Coursework Patterns Project funded by the Office of Educational Research and Improvement (OERI) have shown that the nine item types of the GRE General Test are sensitive to the effects of undergraduate coursework.

All four of the instruments developed since 1987 are designed to assess general education outcomes.

Of the five tests available prior to 1987, only the College Outcome Measures Project (COMP) examination was designed specifically to measure the outcomes of general education programs. Available both as an Objective Test and as a longer Composite Examination, the COMP exam provides a total score, three content subscores, and three process subscores; the Composite Examination also provides several measures of reasoning and communicating. While research by ACT shows that this exam is sensitive to general education program quality, studies at several institutions indicate that the exam is much more sensitive to individual differences and that the effects attributable to individual differences may mask program effects.

All four of the instruments developed since 1987 are designed to assess general education outcomes. The Academic Profile, for example, is intended to measure four academic skills across three content areas. Results of pilot testing during the 1987-88 academic year indicated extremely high inter-correlations among the subscales. As a result, a new version of this test has been prepared for administration during the 1988-89 academic year.

The Collegiate Assessment of Academic Proficiency (CAAP) is intended to measure skills typically attained during the first two years of college. This test consists of four modules that can be administered individually or in combination. In addition, a writing sample based on two independent prompts can be substituted for the multiple choice writing skills module. Because this instrument is just beginning its pilot testing phase, information on reliability and validity is not available.

In the spring of 1988, the College Board introduced the Education Assessment Series (EAS) examinations. Designed to measure learning during the first two years of college, the EAS consists of two tests: one in English composition and one in mathematics. Although the EAS tests are in the developmental stage, national norms are being developed to

Assessment Update
Spring 1989
Volume 1, Number 1

allow institutions to compare the performance of their students with the performance of students across the country.

It is worth noting that the eight examinations described above are norm-referenced. In contrast, the College Basic Academic Subjects Examination (College BASE) is a criterion-referenced achievement test. In addition, national norms will be offered for comparative purposes. College BASE is available in a variety of forms, each with different skill reports. At a minimum, scores will be available in several subject areas, and each subject area will have competency and skill subscores. To date, information on the reliability and validity of College BASE is not available.

In sum, at least nine tests are available for use as general education assessment instruments. Unfortunately, the widespread availability of these measures has not been paralleled by extensive research on their reliability and validity as assessment tools. Institutions interested in assessing general education outcomes using one or more of these measures should proceed cautiously, first determining the extent to which the test matches general education curricula, then empirically evaluating the reliability and validity of test scores.

Addresses of test publishers include the following: *Academic Profile*: ETS College and University Programs, Educational Testing Service, Princeton, NJ 08541; *ACT Assessment Program, College Outcome Measures Project*, and *Collegiate Assessment of Academic Proficiency*: American College Testing Program, P.O. Box 168, Iowa City, IA 52240; *College BASE*: The Riverside Publishing Company, 8420 Bryn Mawr Avenue, Chicago, IL 60631-3476; *College-Level Examination Program* and *Education Assessment Series*: CLEP Program Director, P.O. Box 6602, Princeton, NJ 08540-9885; *Graduate Record Examinations*: Graduate Record Examinations, Educational Testing Service, Princeton, NJ 085481-6000; *Scholastic Aptitude Test*: Scholastic Aptitude Test, Educational Testing Services, Princeton, NJ 085481-6000.

Institutions interested in assessing general education outcomes using one or more of these measures should proceed cautiously.

Criteria for Selecting an Assessment Instrument

When colleges and universities undertake validity studies, the focus of their investigations is score use*. . . . Their primary concern must be with the appropriateness and consequences of* actions *taken on the basis of test data.*

Assessment Update
Winter 1989
Volume 1, Number 4

Washington's Joint Task Force has reported its efforts to evaluate three tests in general education, and similar studies have been conducted at the University of Tennessee. Unfortunately, these studies have been conducted on an ad hoc basis and have not provided a set of standards or a methodology for evaluating assessment instruments.

My purpose in this column is to suggest one possible set of standards for evaluating the appropriateness of using achievement tests as assessment instruments. These standards can be applied to tests in general education or in the major, whether they are commercially available or locally developed.

I propose the use of the concept of validity to evaluate assessment instruments. *Validity* refers to the accuracy and appropriateness of test scores and their interpretations and uses. Traditionally, validity has been divided into three types: content, construct, and criterion-related. Today, many measurement scholars, such as Samuel Messick, argue that construct validity provides an all-inclusive methodology for examining score interpretation and use.

When colleges and universities undertake validity studies, the focus of their investigations is score *use*. Judgments about the validity of score meanings are important, but postsecondary institutions are first and foremost users of test information. Their primary concern must be with the appropriateness and consequences of *actions* taken on the basis of test data.

Before any judgments about the validity of score use can be made, colleges and universities must define the constructs that test scores should represent. Such concepts as effectiveness and quality must be defined in relationship to the mission and goals of the institution. Since outcomes to be assessed are inherent in the mission and goals, the constructs represented by test scores will, to some extent, be peculiar to each institution.

Once constructs have been identified, a variety of approaches can be used to evaluate their relationships to score use. Three criteria that I have used to

BEST COPY AVAILABLE

evaluate the validity of assessment instruments are the substantive, structural, and external components of construct validity, as described by Jane Loevinger.

The *substantive* component of construct validity focuses on how well test items are accounted for by a construct. Content representativeness and dependability of measurement are important aspects of this component. Samuel Messick notes that a trade-off frequently occurs between content representativeness and dependability of measurement because attempts to faithfully represent all aspects of a construct can result in inclusion of items with large errors of measurement. Conversely, attempts to develop highly reliable tests frequently result in omission of important content areas.

The *structural* component of construct validity focuses on the extent to which relationships among test items accurately reflect the structure of the construct. At the item level, this component deals with the approriateness of the scoring model used to represent the construct. At a more general level, the focus is on whether relationships among subscales are consistent with the assumed structure of the construct.

Three criteria that I have used to evaluate the validity of assessment instruments are the substantive, structural, and external components of construct validity.

Research related to the *external* component of construct validity focuses on the extent to which relationships between test scores and other measures are consistent with theories of the construct. For example, measures of program quality should be related to measures of educational experience (coursework and involvement) and relatively insensitive to factors beyond the control of colleges and universities (background characteristics).

Although a variety of research techniques can be used to evaluate these three components of construct validity, I have found that comparing test content to an institution's curriculum and goals provides an excellent test of content representativeness. Because there is a potential tradeoff between content coverage and dependability, it is advisable to assess the reliability or generalizability of test scores. Factor-analysis procedures (either exploratory or confirmatory) provide a good test of the structural component of construct validity, and structural equations can be used to assess the relationships among scores and external variables.

Suggested Readings

Feldt, L. S., and Brennan, R. L. "Reliability." In R. L. Linn (ed.), *Educational Measurement*. (3rd ed., pp. 105-146). New York: Macmillan, 1989.

Loevinger, J. "Objective Tests as Instruments of Psychological Theory." *Psychological Reports*, 1957, 3, 635-694.

Messick, S. "Validity." In R. L. Linn (ed.), *Educational Measurement*. (3rd ed., pp. 13-104). New York: Macmillan, 1989.

Assessment Update
Winter 1989
Volume 1, Number 4

Joint Task Force of the Interinstitutional Committee of Academic Officers and State Board for Community College Education

The results of a pilot study in Washington State strongly suggest that the Academic Profile, the COMP, and the CAAP are not appropriate or useful for assessing sophomores' skills.

The Washington State institutions of higher education have a long-standing commitment to assessing student learning and discerning the value of a college education. They agree that assessment helps enhance the quality of programs. Faculty and administrators across the state are currently involved in discussions and studies to determine the best methods for obtaining appropriate and useful information from assessment activities.

Some states use standardized tests to measure students' academic performance. The state Higher Education Coordinating (HEC) Board recommended in its master plan (December 1987) that two-year and four-year institutions conduct a pilot study to evaluate the appropriateness of using standardized tests as one means for measuring the communication, computation, and critical-thinking skills of sophomores. The sophomores of such a test program would be for institutions to strengthen their curricula, improve teaching and learning, and provide accountability to the public.

To design and implement the study requested by the master plan, two task forces were established. One represented the public baccalaureate institutions, and the other represented the community colleges. Both task forces included faculty and academic administrators from each participating institution, as well as two HEC Board members. The two task forces worked in parallel and ultimately conducted a joint study.

Only three tests met the criteria of the HEC Board's recommendation for study: the Academic Profile, the College Outcome Measures Program (COMP), and the Collegiate Assessment of Academic Proficiency (CAAP). Over 1,300 sophomores from public four-year institutions and from eight two-year colleges were tested, and each student took two of the three tests. More than 100 faculty members from the same institutions took shortened versions of the tests and critiqued them for appropriateness of content and usefulness.

Assessment Update
Fall 1989
Volume 1, Number 3

BEST COPY AVAILABLE

The results of the pilot study strongly suggest that the three tests do not provide an appropriate or useful assessment of the communication, computation, and critical-thinking skills of Washington college sophomores. None of the tests measured the separate academic skills (communication, computation, and critical thinking); rather, these tests primarily measured verbal and quantitative aptitude. Moreover, the tests added little reliable new information about students' academic performance. Results essentially reiterated what is already known from admissions test data and grades. Further, test scores were not sensitive to specific aspects of the college experience, such as estimated time spent studying and credits earned. Finally, none of the tests was judged by faculty as providing an adequate match with curricular content or as being an appropriate or useful measure of communication, computation, and critical thinking.

Norms for making comparisons with peer institutions are currently unavailable. Furthermore, student performance is affected by differences in how institutions administer tests, in the timing of tests, in the selection of students, and in student motivation. Thus, comparisons with future norms which are based on tests given under differing conditions will be misleading.

Student performance is affected by differences in how institutions administer tests, in the timing of tests, in the selection of students, and in student motivation.

Analyses of the costs associated with conducting the pilot study suggests that the projected expense of statewide implementation (testing either a sample of sophomores or all sophomores) would be high and would probably exceed the value of the results.

Both two-year and four-year faculty participants in the study recognized the importance and value of having public as well as institutional access to appropriate measures of student performance. They reaffirmed the value of assessment activities for strengthening the curriculum, improving teaching and learning, and enhancing overall instructional quality. They also shared the view that the development of meaningful assessment measures is both difficult and time-consuming, that measures should be institution-specific, and that national standardized multiple-choice tests have serious limitations for the assessment of teaching and learning.

Copies of the 14-page general report of the task force can be obtained from Dr. Robert M. Thorndike, Department of Psychology, Western Washington University, Bellingham, WA 98225. Copies of the technical report can be obtained from the same source for $10 each.

The College Outcome Measures Program Objective Test

The COMP exam is most appropriate for evaluating general education programs designed to foster the development of generic higher order cognitive processes.

In 1976 the American College Testing (ACT) Program organized the College Outcome Measures Program (COMP) to develop a measure of "knowledge and skills relevant to successful functioning in adult society" (Forrest, 1982, p. 11). Available since 1979-80, the COMP exam has been administered at least once at more than 500 colleges, and it is used annually by approximately 100 four-year institutions in the evaluation of their general education programs. COMP staff acknowledge that the examination is not appropriate for all institutions. They indicate that the COMP exam is most appropriate for evaluating general education programs designed to foster the development of generic higher order cognitive processes, as opposed to discipline-specific or content-specific outcomes (Steele, 1991a).

Since 1983, the COMP exam has been used in Tennessee to evaluate the general education programs of state colleges and universities and to award millions of dollars in public funds for higher education. Until recently, the COMP exam was the only instrument designed for evaluating general education programs, and in 1992, it remains the only measure for which a substantial amount of data is available.

The COMP exam is available in two forms: the Objective Test (consisting of multiple-choice items) and the Composite Examination (containing multiple-choice questions along with exercises requiring students to write essays and record speeches). ACT reports that the correlation between scores on the Objective Test and Composite Examination is .80, allowing the Objective Test to serve as a proxy for the Composite Examination (Forrest and Steele, 1982). COMP staff recommend that the Composite Examination be used to evaluate the performance of individuals and that the Objective Test be used to evaluate general education programs (Steele, 1991c). Most institutions, including the public colleges and universities in Tennessee, use the Objective Test for program evaluation because it is easier to administer and score.

The Objective Test takes approximately 2-1/2 hours to administer and contains 60 questions, each with two correct answers. The questions are

Assessment Update
January-February 1992
Volume 4, Number 1

BEST COPY AVAILABLE

divided among 15 separately timed activities drawing on material (stimuli) from television programs, radio broadcasts, and print media. Students taking the COMP Objective Test are instructed that there is a penalty for guessing (that is, incorrect answers will be subtracted from their scores), but that leaving an answer blank will not be counted against them (Forrest and Steele, 1982).

The combination of two correct answers for each question, the guessing penalty, and no penalty for not answering a question means that the score range for each question is from -2 to 2 points. A score of -2 represents two incorrect answers, while a score of 1 represents one incorrect answer and one answer left blank. A score of 0 represents either both answers left blank or one correct and one incorrect answer. A score of 1 represents one correct answer and one answer left blank, while a score of 2 represents two correct answers. To improve the interpretability of scores, they are rescaled (0-4), making the maximum possible score on the Objective Test 240 points and a chance score 120 points.

Many colleges and universities are drawn to the COMP exam because it provides objective evidence of student intellectual growth over the course of a college education.

New forms of the Objective Test are developed on an annual basis. To ensure the comparability of scores across forms, the COMP staff equates each new form to the original test (Form III). This equating is performed using samples of high school and college seniors who are double-tested using the new form and a previous form of the Objective Test. Statistical procedures involve the use of Angoff's Design II method (J. M. Steele, personal communication, September 14, 1989).

In addition to a total score, the COMP Objective Test provides three content subscores (Functioning Within Social Institutions, Using Science and Technology, and Using the Arts) and three process subscores (Communicating, Solving Problems, and Clarifying Values). In the technical manual for the COMP exam, ACT staff report that the alpha reliability (internal consistency) of the total score is .84 for individuals, and that alpha reliability estimates for the subscores range from .63 to .68 (Forrest and Steele, 1982). Estimates of parallel-forms reliability are .79 for the COMP total score and range from .53 to .68 for the subscores. Most recently, the COMP staff have reported that the generalizability (reliability) coefficients for institution means exceed .90 (for total score and subscores) when the means are based on samples of at least 200 students (Steele, 1989; 1991b).

Many colleges and universities are drawn to the COMP exam because it provides objective evidence of student intellectual growth (value-added) over the course of a college education. Students who persist at an institution can be tested on entrance and again at the end of two or four years of college to determine the growth attributable to their educational experiences (Forrest, 1982).

Partly because many institutions are unwilling to wait two or four years to evaluate student learning, COMP staff provide an estimate of student gain. Based on the fact that the correlation between freshman total scores on the Objective Test and entering ACT Assessment composite scores is .70, the

COMP staff have constructed a concordance table from which institutions may estimate mean freshman COMP scores based on ACT Assessment composite scores. By subtracting estimated freshman score means from actual senior COMP score means, an estimate of score gain (value-added) can be obtained.

References

Forrest, A. *Increasing Student Competence and Persistence: The Best Case for General Education.* Iowa City, Iowa: ACT National Center for the Advancement of Educational Practice, 1982.

Forrest, A., and Steele, J. M. *Defining and Measuring General Education Knowledge and Skills.* Iowa City, Iowa: American College Testing Program, 1982.

Steele, J. M. *College Outcome Measures Program (COMP): A Generalizability Analysis of the COMP Objective Test (Form 9).* Unpublished manuscript, College Outcome Measures Program, Iowa City, Iowa.

Steele, J. M. *Assessing General Education Outcomes of College.* Unpublished manuscript, College Outcome Measures Program, Iowa City, Iowa, 1991a.

Steele, J. M. *College Outcome Measures Program (COMP): A Generalizability Analysis of the COMP Objective Test (Form 10).* Unpublished manuscript, College Outcome Measures Program, Iowa City, Iowa, 1991b.

Steele, J. M. *Evidence for What? Fitting Assessment to the User.* Unpublished manuscript, College Outcome Measures Program, Iowa City, Iowa, 1991c.

For additional information contact Dr. Joe M. Steele, College Outcome Measures Program, P.O. Box 168, Iowa City, IA 52243. Tel: (319) 337-1121.

GARY R. PIKE

Academic Profile II

During 1987-88 the Educational Testing Service (ETS) introduced the *Academic Profile*, a test of academic skills designed to assist colleges and universities in assessing the quality and effectiveness of their general education programs. During the first two years of pilot testing, the *Academic Profile* was used by more than 200 two- and four-year institutions to test over 60,000 students.

> ***ETS introduced the*** **Academic Profile,** ***a test of academic skills designed to assist colleges and universities in assessing the quality and effectiveness of their general education programs.***

Assessment Update
September-October 1991
Volume 3, Number 5

The original version of the *Academic Profile* was available in both short and long forms. The long form provided reliable scores for both individuals and groups, and measured four types of skills (reading, writing, critical thinking, and mathematics) in three content areas (humanities, social sciences, and natural sciences). Highly reliable total scores were also provided for both individuals and groups. The short form of the exam provided reliable group scores for the four skill areas and the three content areas, as well as reliable total scores for both individuals and groups. An optional written essay was available with both forms of the exam.

Based on two years of pilot testing, ETS identified several problems with the original version of the *Academic Profile*, as well as several opportunities for improved score reporting. In 1989-90 ETS began pilot testing a new version of the test, the *Academic Profile II*. Changes in the test included significantly lower intercorrelations among the four skills dimensions and a 40-minute short form that could be administered in a standard class period. Also, ETS began using scaled scores for total score and the seven subscores.

Related to the first change, factor analyses identified three factors as being measured in the skill areas: writing, mathematics, and reading/critical thinking. The factor structure of the *Academic Profile II* allowed development of several new performance-level (criterion-referenced) indicators. For the long form of the exam, seven performance indicators are provided: writing, mathematics, reading/critical thinking (combined), reading/critical thinking (humanities), reading/critical thinking (social sciences), and reading/critical thinking (natural sciences). Three performance indicators

BEST COPY AVAILABLE

are available to users of the short form: writing, mathematics, and reading/critical thinking (combined). Three levels of performance are described for each indicator.

To assist institutions in interpreting score information, two types of comparative (norm-referenced) data are provided. First, *Standard Reference Group* reports that aggregate the mean scores of all institutions testing at least 30 students in a given year are provided for total score, the four skill subscores, and the three content subscores. Second, institutions may select a group of peers from a list of user institutions and *Self-Selected Reference Group* reports are provided for total score and subscores.

Institutions may select a group of peers from a list of user institutions and Self-Selected Reference Group *reports are provided for total score and subscores.*

In addition to the multiple-choice questions on the long and short forms of the new *Academic Profile II*, several demographic questions are included and a 45-minute essay is available on an optional basis. Institutions can add up to 50 locally developed multiple-choice items to the test.

The technical manual for the *Academic Profile II* indicates that the reliability of total score on the long form for individuals is .94, and reliability estimates for the skills and content subscores range from .74 (critical thinking) to .85 (reading). Reliability of total score on the short form for individuals is .80. Reliability estimates for institutional mean scores on the long form are .99 for total score and all subscores. Reliability estimates for institutional means on the short form are somewhat lower. They range from .88 (social sciences) to .95 (mathematics). The group mean reliability coefficient for total score on the short form is .90.

The technical manual also reports the results of three analyses designed to evaluate the construct validity of total scores and the seven norm-referenced subscores on the *Academic Profile II*. Specifically, the research indicates that total scores and subscores are significantly higher for students with a cumulative grade point average above 3.00 (on a 4-point scale) than for students with grade point averages below 3.00. Research on score differences by class rank (freshman, sophomore, junior, and senior) also generally supported the validity of scores on the *Academic Profile II*. In some cases, however, means for adjacent categories (such as sophomore vs. junior) were not significantly different. Relationships between skill and content scores by proportion of the core curriculum completed also indicated that these scores are positively related to quantity of coursework.

As previously noted, factor analysis of item-type scores from the *Academic Profile II* indicates the existence of three underlying factors corresponding to writing, mathematics, and reading/critical thinking. The results are presented by ETS as evidence of the construct validity of the criterion-referenced proficiency level scores on the test. The test developers also indicate that slightly over two-thirds of the students they sampled showed differentiation in their proficiency scores. That is, two-thirds of the students had different levels of performance depending on the proficiency score being considered. The authors of the technical manual conclude that these results provide evidence of the discriminant validity of the proficiency scores.

Assessment Update
September-October 1991
Volume 3, Number 5

Readers interested in more information about the *Academic Profile* may contact Nancy Beck, ETS College and University Programs, Princeton, NJ 08541. Tel: (609) 243-8195.

During the 1990-91 academic year, the University of Tennessee, Knoxville, has been engaged in a project to evaluate the *Academic Profile II*. Results of this research, *The Evaluation of the Academic Profile*, can be obtained from the Center for Assessment Research and Development, The University of Tennessee, Knoxville, 1819 Andy Holt Avenue, Knoxville, TN 37996-1350.

References

College Board and Educational Testing Service, *The Academic Profile: User's Guide*. Princeton, NJ: College Board and Educational Testing Service, 1990.

College Board and Educational Testing Service. *Higher Education Assessment Newsletter*. Princeton, NJ: College Board and Educational Testing Service, 1990.

College BASE

The College Basic Academic Subjects Examination (College BASE) is one of the newest standardized tests designed for the assessment of general education outcomes. Unlike other tests of this genre, such as the COMP and CAAP exams, the College BASE is a criterion-referenced achievement test that focuses on the degree to which students have mastered particular skills and competencies commensurate with the completion of general education coursework. College BASE is appropriate for assessing either students or programs. Consequently, the test has been used to evaluate academic programs at a variety of institutions, and it has been used as a selection criterion for entry into teacher education in Missouri.

The College BASE is a criterion-referenced achievement test that focuses on the degree to which students have mastered particular skills and competencies.

College BASE assesses achievement in four subject areas: English, mathematics, science, and social studies. Subject area scores are built on content clusters, which in turn are based on enabling skills. For example, mathematics scores are based on three content clusters: general mathematics, algebra, and geometry. The cluster score for geometry is derived from skills related to the recognition of two- and three-dimensional figures and the ability to perform geometric calculations. Numerical scores are provided for each subject area and cluster, along with a composite (total) score. Ratings of "high," "medium," or "low" are provided for each enabling skill.

In addition to the evaluation of content areas, the College BASE assesses three competencies across disciplines: interpretive reasoning, strategic reasoning, and adaptive reasoning. According to the test's authors, interpretive reasoning is the most basic level of information processing, beyond factual recall, and includes such abilities as paraphrasing, summarizing, and explaining. In contrast, strategic reasoning includes skills related to definition, comparison, and classification. Adaptive reasoning includes skills related to definition, comparison, and classification. Adaptive reasoning includes the skills of synthesis and evaluation.

Assessment Update
January-February 1991
Volume 3, Number 1

The College BASE is available in three forms. The long form includes all four content areas and takes approximately 2-1/2 hours to administer. The short form includes only English and mathematics and can be administered in one hour and twenty minutes. The third form is an institutional-matrix form and takes approximately 40 minutes per student to administer. Written essays may be included with all three forms. The long and short forms of the exam are appropriate for both individual and institutional assessment, while the institutional-matrix form provides only institutional scores.

Item responses on the College BASE are calibrated and scaled with a two-parameter logistical item-response model. The scale for the composite (total) score, as well as for the subject, cluster, and competency scores, has a mean of 300 and a range of from 40 to 560 points. The standard deviation for each score is set at 65 points. Students' performance on the test closely parallels its specifications. For example, a score of 299 represents the 50th percentile on the test, compared to a theoretical mean of 300 points.

The authors of the test report that KR-20 estimates of the reliability (internal consistency) of scores on the cross-disciplinary College BASE range from .77 (English) to .89 (mathematics) for the subject area scores. These estimates for the cluster scores range from .73 (adaptive reasoning) to .868 (strategic reasoning).

A recent study of the dependability of group means on the College BASE conducted at the University of Tennessee, Knoxville (UTK) found that institutional scores on the test are quite dependable, even when they are based on relatively small samples of students. For example, samples of 50 students' generalizability coefficients for all four subject area scores are in excess of .90.

As evidence of the validity of the College BASE, its authors have conducted factor analysis of test scores that have reproduced the subscales specified on the test. Analyses of College BASE scores also indicate that the test has significant positive correlations with students' scores on the ACT Assessment Examination, Scholastic Achievement Test (SAT), and cumulative grade point average, the authors say. Research on the administration of the test to more than 1,000 students at UTK provides additional evidence of test validity. First, factor analysis of the nine cluster scores on the long form of the test reproduced the test's four subject area scores. Second, analysis of the relationships between students' subject area scores and their coursework in history, humanities, mathematics, natural science, and social science revealed that coursework in mathematics was significantly related to mathematics scores, natural science coursework to science scores, and history coursework to social studies scores. Counter to expectations, humanities coursework was not significantly related to English scores.

Our experiences with the College BASE indicate that this test is the most appropriate standardized test we have investigated for evaluating the effects of our general education program.

Overall, our experiences with the College BASE at UTK indicate that this test is the most appropriate standardized test we have investigated for evaluating the effects of our general education program. Its strengths lie in the number of specific subscores provided, the evidence of the construct validity of those subscores, and the relationship of content areas to specific patterns of coursework. Despite these strengths, an evaluation of the test content by a faculty committee revealed that the test covers only about 36% of the general education goals at UTK. Thus, while this test may be the most appropriate standardized test for UTK, it suffers the same limitation as other standardized tests do in that it covers only a fraction of UTK's general education outcomes. Faculty at any other institution considering the College BASE would need to compare its stated outcomes for general education with the specifications developed for the exam, in order to determine the appropriateness of the College BASE for its own purposes. It is generally accepted, however, that any standardized test used for assessment should be supplemented by other measures, such as essays and performance appraisals.

As a concluding note, it should be emphasized that the results for UTK are specific to our institution and should not be taken as an indication of the validity of the College BASE at any other institution. Nevertheless, the results do suggest that institutions interested in using standardized tests to assess general education outcomes should give serious consideration to the College BASE.

Any standardized test used for assessment should be supplemented by other measures, such as essays and performance appraisals.

For Further Reference

Center for Educational Assessment, University of Missouri-Columbia. *College BASE*. Chicago, Riverside Publishing, 1988.

Osterlind, S. J. *College BASE: Guide to Test Content*. Chicago: Riverside Publishing, 1989.

Osterlind, S. J. *College BASE: Technical Manual*. Chicago: Riverside Publishing, 1990.

Pike, G. R. "The Components of Construct Validity: A Comparison of Two Measures of General Education Outcomes." *Journal of General Education*, 1992, *41*, 130-139.

Assessment Update
January-February 1991
Volume 3, Number 1

GARY R. PIKE

Using Self-Report Measures as Proxies for Test Scores

Despite the increased use of outcomes assessment for program evaluation, criticism of postsecondary education continues. Last year, for example, the Wingspread Group on Higher Education issued a report criticizing higher education and calling on colleges and universities to redouble their assessment efforts. The recently enacted National Education Goals represent another effort to promote assessment of student outcomes. Specifically, National Education Goal 5.5 states that by the year 2000, "The proportion of college graduates who demonstrate an advanced ability to think critically, communicate effectively, and solve problems will increase substantially." In order to monitor progress toward Goal 5.5, the U.S. Department of Education has proposed that a test similar to the National Assessment of Educational Progress be developed and administered to college students.

Despite the increased use of outcomes assessment for program evaluation, criticism of postsecondary education continues.

Several participants in a federally sponsored workshop voiced reservations about the feasibility of developing and implementing such a test. Trudy Banta, for example, raised questions about the desirability and practicality of achieving national consensus on the outcomes to be assessed, while Stephen Dunbar identified several potential technical problems in creating a test that would be reliable and valid for a national assessment. Other participants were more optimistic. James Ratcliff concluded that a national assessment program was possible. However, he urged that the development of a national assessment of college student learning be a long-term goal. In the interim, Ratcliff argued, alternative measures should be used as proxies for a national test of college outcomes.

In 1991, the Resource Group on Adult Literacy and Lifelong Learning of the National Education Goals Panel also recommended that alternatives to a national test be considered. The resource group argued that measures of good practice in postsecondary education could serve as proxies for a national test. These measures of good practice could also be used to collect data that would suggest policy options for improving student achievement.

Assessment Update
July-August 1994
Volume 6, Number 4

BEST COPY AVAILABLE

Based on the results of the workshop and the resource group's report, the National Center for Education Statistics contracted with the National Center for Higher Education Management Systems (NCHEMS) to conduct a preliminary study concerning the feasibility of using measures of good practice as indicators of the quality and effectiveness of postsecondary education. NCHEMS staff reviewed previous research on a variety of possible indicators, including institutions' general education requirements, reliance on active learning practices in teaching, and students' reports of their involvement or cognitive development, or both, during college. The NCHEMS report concluded that students' reports of active learning, involvement, and cognitive development during college all had moderate to high potential as indicators of good educational practice.

I became interested in this topic after reading the NCHEMS report on measures of good practice. First in Tennessee, and then in Missouri, I undertook two research projects to evaluate the use of students' reports of their college involvement and the cognitive gains they made during college as alternatives to standardized tests of student achievement. I want to share the preliminary results of these studies because I believe that they have important implications for combining self-report (survey) data with scores from standardized achievement tests.

Measures of students' extracurricular academic involvement and faculty-student interaction had the greatest positive effects on achievement test scores.

The standardized test that I used in both studies was the College Basic Academic Subjects Examination (College BASE). A description of this test is available in an earlier *Assessment Measures* column (*Assessment Update*, 1991, Vol. 3, No. 1). In the first study, the measures of involvement used were modeled after the quality-of-effort scales on the College Student Experiences Questionnaire (CSEQ) (see this column, *Assessment Update*, 1990, Vol. 2, No. 1), as well as scales on faculty-student interaction developed by Pascarella and Terenzini. Self-report measures of cognitive development in the first study were drawn from items on the alumni survey used in Tennessee.

The second study focused exclusively on the relationship between self-reports of cognitive development and standardized test scores. The measures of cognitive development that I used in the second study were drawn from the content specifications for College BASE. Specifically, I took 28 of the enabling subskills from College BASE and transformed them into survey questions. For example, one of the questions in the self-assessment survey asks students to rate their ability to determine language meaning in passages that use connotation and figures of speech, one subskill that serves as a foundation for the English section of College BASE.

Three general conclusions emerged from these studies. First, the results of the first study using College BASE scores and CSEQ-like involvement questions revealed that there are moderate to strong relationships between students' self-reports of their involvement during college and scores on College BASE. In this study, measures of students' extracurricular academic involvement and faculty-student interaction had the greatest positive effects on achievement test scores. In fact, as much as one-quarter of the variability

Assessment Update
July-August 1994
Volume 6, Number 4

in students' scores could be explained by differences in their levels of involvement during college.

The second conclusion, confirmed in both studies, suggests that there is a weak to moderate relationship between self-reports of cognitive development and achievement test scores. One reason these relationships were only weak to moderate is that the different methods used to measure constructs, such as English or mathematics ability, influence how students respond. The end result is that it is not sufficient to look at the simple relationship (such as correlations) between measures based on different methods. Use of self-reports as proxies for test scores is possible, but it requires the use of sophisticated statistical techniques that can remove the method-specific influences from assessment data.

The third conclusion to emerge from the two studies is that the strength of the relationship between self-reports of cognitive development and test scores is directly related to the correspondence between items on the two measures. Comparing results across the two studies, the relationship between self-reports from the CSEQ and College BASE scores was much weaker than the relationship between self reports derived from the College BASE test specifications and College BASE scores.

Achievement tests can tell assessment researchers how *students perform, but they cannot tell us* why *people perform the way they do.*

Based on the results of my two studies, the answer to the question "Can self-reports be used as proxies for test scores?" is a qualified yes. Test scores and self-reports are related to each other, but the relationship is not simple. The magnitude of the relationship is influenced by a variety of factors, including similarity of content and methods used to elicit assessment data. In addition, my research suggests that the most profitable use of self-reports is not as proxies for test scores but rather as explanatory variables. As several writers have observed, achievement tests can tell assessment researchers *how* students perform, but they cannot tell us *why* people perform the way they do. Answering the important question of why requires a careful integration of objective measures of student achievement and measures of students' college experiences.

Assessment Update
July-August 1994
Volume 6, Number 4

Measuring Growth in Student Learning and Development

While there is virtually unanimous agreement that student change during college should be assessed, there is little agreement about how it should be assessed.

Assessment Update
March-April 1994
Volume 6, Number 2

I must confess that this is not the column I had planned to write. About a week ago, I was contacted by a state higher education commission about the feasibility of developing measures of "value added" for *College BASE*. Although this topic has been the subject of one *Assessment Measures* column and rejoinder and response columns (see *Assessment Update*, 1992, Vol. 4, No. 2), I feel compelled to return to the issue of assessing student change during college. My purpose is to acquaint readers with some of the options available to institutions interested in studying student change and to provide a set of references on the measurement of change.

The rationale for studying student change is convincing. Numerous authors argue that change is a fundamental feature of higher education and must be studied. These individuals also argue that studies of how students change during their college careers provide a more accurate representation of educational effects than do traditional outcomes studies because they eliminate the confounding effects of differences in ability levels. Other often-cited advantages of studying change include encouraging faculty and administrators to think in developmental terms, promoting greater faculty and student involvement in assessment, and providing outcomes measures that are appropriate for nontraditional programs and students.

While there is virtually unanimous agreement that student change during college should be assessed, there is little agreement about *how* it should be assessed. In the remainder of this column, I will review several methods of representing change during college. Readers who hope I will identify the "right" way of studying change will be disappointed; I firmly believe that there is no right way. Furthermore, the method that may be most useful for one campus may not be the most useful method for another campus.

Gain Scores. One of the most widely publicized and used methods of studying change during college is the gain score. Until recently, this was a criterion for awarding state funding supplements in Tennessee. Gain scores are regularly reported by the College Outcome Measures Program (COMP)

BEST COPY AVAILABLE

staff as indicators of student learning and program effectiveness. Calculating a gain score entails administering an instrument to students at the beginning of a program of study and then readministering the instrument on completion of the program. The difference between the two scores is a measure of student growth, and the average of all students' scores is a measure of institution, or program, effectiveness. Steele (1989) has provided a detailed discussion of how gain scores can be used for program evaluation.

The articles by Steele and myself in the previous *Assessment Update* provide an indication of the controversial nature of gain scores. In addition to that discussion, those interested in considering gain scores may wish to consult the articles by Baird (1988), Linn (1981), and Pike (1992a).

Estimated Gain Scores. One limitation of using gain scores is that assessment researchers must wait at least two or four years before they have any data on program outcomes. Partly in response to the long timelines in studies of student gain, the developers of the COMP exam have offered to provide an *estimate* of student gain in total score. Using the fact that the correlation between ACT Assessment composite scores and freshman COMP scores is approximately 0.70, ACT has developed concordance tables that allow an institution to estimate mean freshman COMP scores for a group of graduating students who have valid ACT Assessment scores. The mean estimated freshman score for an institution is then subtracted from the mean score for the same cohort of graduating students to provide an estimate of gain.

The use of estimated score gain in outcomes assessment has been discussed by Steele (1989) and Banta and others (1987). Although estimated gain eliminates the need for longitudinal studies covering several years, institutions should not adopt this method without some careful research. Assessment professionals need to determine that the estimates for freshman scores are appropriate for their campuses. Estimates derived from a national data set may not be appropriate for a specific campus. Further, as an institution adopts a new form of a test, the estimates of freshman scores on that test must be checked to see whether they are still valid. Another potential limitation of the method is that entering ACT Assessment scores may not be available for some graduating students. Banta and others (1987) compared graduating seniors who did and did not have ACT Assessment scores and found that the two groups differed in race, socioeconomic status, high school grade point average, and college experiences.

Although estimated gain eliminates the need for longitudinal studies covering several years, institutions should not adopt this method without some careful research.

Residual Scores. In his most recent book, Astin (1993) advocates the use of residual scores in assessing the net effects of college. Residual scores are calculated by regressing students' scores at the end of a program of study on one or more entry measures in order to develop a prediction model. According to Astin, the differences between actual and predicted scores represent the net effects of college experiences. This approach is being used by the National Center on Postsecondary Teaching, Learning and Assessment to study college effects (Pascarella, personal communication). Technical discussions of the strengths and limitations of residual scores can be found in research by Astin (1993), Baird (1988), and Pike (1992a).

An approach related to the use of residual scores, the use of fixed-effect structural equation models to study change, is discussed by Pike (1991). A primary limitation of this method is that it is very demanding technically (statistically).

Mixed-Effect Models. All of the approaches discussed thus far seek to assess student achievement at the end of a program of study after *controlling* for differences in entering achievement (for example, freshman test scores). These approaches represent a general class of statistical analyses that can be termed fixed-effect models. Recently, authors such as Muthén (1991) have suggested approaches based on mixed- or random-effects models. Instead of seeking to eliminate individual differences using statistical procedures, random-effects models explicitly consider individual differences in change. McLean and Sanders at the University of Tennessee, Knoxville, have proposed a method of using mixed models for teacher evaluation that deserves additional study. Readers interested in learning more about that approach should contact the authors directly. The mixed-effects approach advocated by McLean and Sanders makes use of measured variables. Other authors have suggested that mixed-effects models be used with latent (unmeasured) variables to study student change (Knight, 1993; Muthén, 1991; Pike, 1992b).

Assessing how students change during college is an important undertaking, but it is not easy.

Growth Curves. If institutions are willing to invest the time and resources to gather data about student learning at more than two points in time, a whole new set of methods for studying student change becomes available. This class of methods involves modeling student growth or learning curves at multiple points in time. Discussions of these approaches are available in the works of Bryk and Raudenbush (1992) and Willett (1988).

As I stated at the outset, assessing how students change during college is an important undertaking, but it is not easy. The difficulty is apparent in the fact that while almost 90% of all colleges and universities are implementing assessment programs, only about 10% to 15% are attempting to study how students change during their college careers. I strongly encourage assessment professionals to examine the various methods available for studying change, to identify methods that might be used by their institution, and then to evaluate carefully the appropriateness of using those methods in their own settings.

For additional information about the McLean/Sanders approach, write William L. Sanders or Robert A. McLean, College of Education, Claxton Education Building, University of Tennessee, Knoxville, Knoxville, TN 37996-3400.

Assessment Update
March-April 1994
Volume 6, Number 2

References

Astin, A. W. *What Matters in College?* San Francisco: Jossey-Bass, 1993.

Baird, L. L. "Value-Added: Using Student Gains as Yardsticks of Learning." In C. Adelman (ed.), *Performance and Judgment: Essays on Principles and Practice in the Assessment of College Student Learning* (pp. 205-216). Washington, D.C.: U.S. Government Printing Office, 1988. Document No. OR88-514.

Banta, T. W., Lambert, E. W., Pike, G. R., Schmidhammer, J. L., and Schneider, J. A. "Estimated Student Score Gain on the ACT COMP Exam: Valid Tool for Institutional Assessment?" *Research in Higher Education*, 1987, *27*, 195-217.

Bryk, A. S., and Raudenbush, S. W. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, Calif.: Sage, 1992.

Knight, W. E. "An Examination of Freshman to Senior General Education Gains Across a National Sample of Institutions with Different General Education Requirements Using a Mixed-Effect Structural Equation Model." *Research in Higher Education*, 1993, *34*, 41-54.

Linn, R. L. "Measuring Pretest-Posttest Performance Changes." In R. A. Berk (ed.), *Educational Evaluation Methodology: The State of the Art* (pp. 84-109). Baltimore: Johns Hopkins University Press, 1981.

Muthén, B. O. "Analysis of Longitudinal Data Using Latent Variable Models with Varying Parameters." In L. M. Collins and J. L. Horn (eds.), *Best Methods for the Analysis of Change* (pp. 1-17). Washington, D.C.: American Psychological Association, 1991.

Pike, G. R. "Using Structural Equation Models with Latent Variables to Study Student Growth and Development." *Research in Higher Education*, 1991, *32*, 499-524.

Pike, G. R. "Lies, Damn Lies, and Statistics Revisited: A Comparison of Three Methods of Representing Change." *Research in Higher Education*, 1992a, *33*, 71-84.

Pike, G. R. "Using Mixed-Effect Structural Equation Models to Study Student Academic Development." *Review of Higher Education*, *15*, 1992b, 151-178.

Steele, J. M. "Evaluating College Programs Using Measures of Student Achievement and Growth." *Educational Evaluation and Policy Analysis*, 1989, *11*, 357-375.

Willett, J. B. "Questions and Answers in the Measurement of Change." In E. Rothkopf (ed.), *Review of Research in Education* (pp. 345-422). Washington, D.C.: American Educational Research Association, 1988.

Freshman-to-Senior Gains at the University of Tennessee, Knoxville

"No institution can have a clear idea of the amount of student growth its general education program may be promoting until it tests its own incoming students, then administers an equivalent form of the same test to graduates."

In one state—Tennessee—the higher education commission bases a percentage of the annual allocation of public funds for higher education instruction on institutional mean scores, mean score gain, and efforts taken to improve students' scores on the ACT-COMP exam. For most institutions in Tennessee, awards for score gain are based on "estimated" gain scores (see my column in the January-February 1992 issue). In 1987, researchers at the University of Tennessee, Knoxville, published an article questioning the validity of using estimated score gain on the COMP exam as an indicator of general education program effectiveness (Banta and others, 1987). The UTK researchers concluded: "No institution can have a clear idea of the amount of student growth its general education program may be promoting until it tests its own incoming students, then administers an equivalent form of the same test to graduates" (p. 216).

As a result of the 1987 findings, a comprehensive study of student growth in college, as measured by actual freshman-to-senior gain scores on the COMP exam, was undertaken at the University of Tennessee, Knoxville, in 1989. The subjects were 942 seniors who had taken the COMP Objective Test first as freshmen, then as seniors. Sources of data included the following: (1) ACT Assessment composite scores, college grades, and transcripts showing courses taken in college, obtained from student records for 873 of the students having freshman and senior COMP scores; (2) responses to the senior survey (designed to gather information about faculty-student interaction, peer interaction, and attendance at campus cultural events), obtained for 722 of the double-tested seniors; and (3) the College Student Experiences Questionnaire (CSEQ), administered to 120 paid volunteers from the double-tested senior sample.

The analysis of freshman-to-senior gains was accomplished in three phases: (1) profiles of students with the highest gain scores (upper quartile) and lowest gain scores (bottom quartile) were compared; (2) an exploration of the relationships between college experiences, as measured by the UTK Senior Survey and the CSEQ, was undertaken; and (3) an evaluation of the

Assessment Update
March-April 1992
Volume 4, Number 2

BEST COPY AVAILABLE

relationship between coursework (obtained from student transcripts) and gain scores was undertaken.

Analysis of high and low gain groups indicated that seniors with the highest gain scores differed most from those with the lowest gain scores in aptitude and achievement. Those with *highest gains* had *lower* ACT Assessment composite scores, ACT English and mathematics scores, and cumulative grade point averages than did seniors with low gain scores. The group with high gain also had lower COMP total scores and subscores as freshmen. The high and low gain groups did not differ on any of the CSEQ quality of effort or involvement scales, nor on any of the scales from the Senior Survey. In terms of coursework, students with low gain scores took fewer business courses and more natural science courses. (Regarding this last finding, it is worth noting that the relationships between coursework and gain were extremely weak.)

Canonical variate analysis was used to explore the relationship between involvement in the college experience (CSEQ and Senior Survey scales) and COMP gain. Data from the CSEQ indicated that investing effort in library experience and experience with faculty produced gains in the categories "Using the Arts" and "Communicating," but spending time in clubs and organizations and in writing activities retarded growth in those categories. Analysis of the Senior Survey responses revealed more contradictory evidence: interacting with peers and attending campus events fostered growth in the category "Clarifying Values," but inhibited gains in "Communicating." Canonical redundancy analysis further indicated that the quality of effort measures of the CSEQ explained only 5% of the variance in subscore gain, while less than 1% of the variance in subscore gain was accounted for by scales from the senior survey.

T*he analysis of freshman-to-senior gains produced a number of disturbing findings.*

Three approaches were used to assess the impact of coursework on gain scores. First, cluster analysis was employed to identify groups of courses that *maximized* score gain on one or more COMP subscales. Four coursework clusters were identified. In general, above average gains on one COMP subscale were offset by below average gains on another. Thus, differences in coursework had virtually no effect on gain in total score.

In the second set of analyses concerning coursework and gain, a three-step process was employed to investigate the relationships between actual patterns of course-taking and gain scores. Multivariate analysis of covariance did not identify a significant effort on COMP gain of coursework actually taken.

In the final set of analyses, data on actual patterns of coursework and gain scores were again evaluated using a latent variable model of change. This analysis confirmed that actual course-taking patterns at UTK had no significant effect on student growth as measured by COMP gain (Pike, 1991).

In summary, the analysis of freshman-to-senior gains produced a number of disturbing findings. The first was the absence of any meaningful relationship between students' investment of time and effort and score gain on the COMP exam. The absence of clear relationships between score gain and college coursework was a second disturbing finding. The significant *negative* relationship between gain and initial (freshman) academic aptitude was perhaps the most disturbing result of the freshman-to-senior gains study because it suggested that the most effective method of improving student score gain on the COMP exam is for an institution to attract less well prepared students.

In the freshman-to-senior gains study, and in my research (Pike, 1992), the UTK researchers examined possible reasons for their counterintuitive findings. They concluded that gain scores are probably not valid measures of student learning and development. The reliability estimate for individual gain scores was found to be unacceptably low (.14) and the 95% confidence interval for mean gain scores was well over 20 points. The UTK researchers also concluded that the negative correlation between gain and freshman COMP scores was probably a spurious relationship produced by measurement error. Moreover, this spurious relationship was found to be responsible for many of the counterintuitive findings. The researchers concluded, as did Lee Cronbach and Lita Furby (1970, p. 80): "It appears that investigators who ask questions regarding gain scores would ordinarily be better advised to frame their questions in other ways."

Gain scores are probably not valid measures of student learning and development.

Copies of *Freshman-to-Senior Gains at the University of Tennessee, Knoxville* are available from the Center for Assessment Research and Development, 1819 Andy Holt Ave., Knoxville, TN 37996-4350. The cost is $10 to cover duplication and postage.

References

Banta, T. W., Lambert, E. W., Pike, G. R., Schmidhammer, J. L., and Schneider, J. A. "Estimated Student Score Gain on the ACT COMP Exam: Valid Tool for Institutional Assessment?" *Research in Higher Education*, 1987, *27*, 195-217.

Cronbach, L. J., and Furby, L. "How Should We Measure 'Change'—Or Should We?" *Psychological Bulletin*, 1970, *74*, 68-80.

Pike, G. R. "Using Structural Equation Models with Latent Variables to Study Student Growth and Development." *Research in Higher Education*, 1991, *32*, 499-524.

Pike, G. R. "Lies, Damn Lies, and Statistics Revisited: A Comparison of Three Methods of Representing Change." *Research in Higher Education*, 1992, *33*.

Assessment Update
March-April 1992
Volume 4, Number 2

GARY R. PIKE

Assessing the Critical Thinking Abilities of College Students

We generally take for granted that some of the most important outcomes of a college education are the abilities to think critically, reason effectively, and solve problems. As Pascarella, Terenzini, and their colleagues have observed, disciplinary knowledge becomes obsolete very quickly. Consequently, one of the most lasting contributions a college or university can make is to cultivate enduring critical thinking skills among its students. Given current interest in critical thinking, it is not surprising that the Campus Trends surveys conducted by the American Council on Education have consistently found that measurement of students' critical thinking skills plays an important role in most campus assessment programs. Improved critical thinking and problem solving are also important components of the National Education Goals and federal efforts to assess college student learning (see Editor's Notes, p. 3).

One of the most lasting contributions a college or university can make is to cultivate enduring critical thinking skills among its students.

Because evaluating students' critical thinking abilities can be an important part of assessing the impact of college on students, I plan to devote several *Assessment Measures* columns to the topic of assessing critical thinking. In this column, I focus on some of the general issues surrounding critical thinking assessments, then I provide a brief survey of the kinds of assessment measures that are available. Subsequent columns will focus on specific measures or classes of instruments.

One of the most basic issues surrounding the assessment of critical thinking is the definition of this term. As Cuban (1984, p. 676) observed, defining what is meant by critical thinking "is troublesome to both social scientists and practitioners. Troublesome is a polite word; the area is a conceptual swamp." I would venture that most of us would say we can define critical thinking. However, I would also wager that our definitions will vary considerably. Many scholars interested in studying critical thinking resort to a laundry list of skills and abilities to define it. For example, Pascarella and Terenzini (1991, p. 118) concluded that critical thinking "typically involves the individual's ability to do some or all of the following: identify central issues and assumptions in an argument, recognize important relationships, make correct inferences from data, deduce conclusions from information

Assessment Update
March-April 1996
Volume 8, Number 2

BEST COPY AVAILABLE

or data provided, interpret whether conclusions are warranted on the basis of the data given, and evaluate evidence or authority." While a listing of critical thinking skills can be helpful, it does not provide a framework from which to view the concept. My own bias is that assessment practitioners need conceptual frameworks (for example, theories of how college affects students) in order to interpret data about student performance and program effectiveness.

From my perspective, one of the more useful conceptual frameworks for viewing critical thinking was proposed by Leonard Baird (1988). Baird's model of college outcomes consists of three interrelated levels: basic skills, general learned abilities, and generic academic outcomes. In Baird's hierarchy, basic skills are fundamental skills that are prerequisite to learning in college. Skills at this level include mastery of the basics of arithmetic and the ability to read and comprehend ordinary paragraphs.

The term for Baird's second level, general learned abilities, undoubtedly comes from his years with the Educational Testing Service (ETS). This term consistently appears in descriptions of what the tests developed by ETS purport to measure. These abilities are broader than basic skills and are presumed to underlie basic academic work. Staff at ETS frequently argue that general learned abilities are largely content-free and represent remarkably stable traits that are not easily affected by education programs.

Assessment practitioners need conceptual frameworks in order to interpret data about student performance and program effectiveness.

At the apex of Baird's hierarchy are generic academic outcomes: learned abilities that develop as a result of college experiences. Although generic academic outcomes are broader than traditional conceptualizations of critical thinking, higher-order thinking skills are a crucial part of these outcomes. I particularly like the fact that Baird's generic outcomes are the product of a student's total college experiences, not a particular course or group of courses. I also find it intriguing that in Baird's model higher-order thinking skills are at the apex of the hierarchy and sensitive to college effects. In most hierarchical models, the focus tends to be on stable traits at the apex. In Baird's model, those stable traits are at the middle level of the hierarchy.

Another issue in the assessment of higher-order thinking skills concerns whether critical thinking is unidimensional or multidimensional. Given that many different types of abilities (for example, interpretation of evidence, integration of information, and evaluation of conclusions) are included within the critical thinking rubric, it seems likely that a multidimensional representation would be preferred. However, many of the most widely used critical thinking measures are unidimensional, as are most documented college outcomes. It may be that while multidimensional models provide a much richer representation of critical thinking, the outcomes associated with college are essentially unidimensional. Robert L. Thorndike (1985, p. 253) reached a similar conclusion when he reviewed the research on the relationships among job performance, training, and general cognitive ability: "In the context of practical prediction, 'g' appears to be alive and well."

Assessment Update
March-April 1996
Volume 8, Number 2

A third issue in assessing critical thinking concerns the nature of student development. For example, should student development be viewed as a continuum in which the attainment of critical thinking abilities is essentially additive, or should development be viewed as a series of stages students pass through as they develop higher-order thinking skills? How this question is answered has a powerful influence on how thinking skills will be measured. When student development is presumed to be continual, traditional objective tests can provide reliable and valid measures of higher-order thinking skills. However, when student development is viewed as a series of discrete stages, traditional objective tests are less useful. In order to provide valid representations of stage development, assessments generally make greater use of expert judgment in assigning individuals to different stages.

How these three issues are addressed can provide a useful heuristic for organizing measures of higher-order thinking skills. For example, one of the most widely used measures of students' critical thinking abilities, the Watson-Glaser Critical Thinking Appraisal, focuses on the types of outcomes described by Pascarella and Terenzini (1991) and is essentially a unidimensional measure of development along a continuum. This test is joined by other standardized objective measures of critical thinking, such as the Cornell Critical Thinking Test and the critical thinking subtest of the Collegiate Assessment of Academic Proficiency.

It may be that while multidimensional models provide a much richer representation of critical thinking, the outcomes associated with college are essentially unidimensional.

While many of the traditionally organized unidimensional measures of critical thinking are objective tests, more subjective measures are not automatically excluded from this category. Performance assessments, such as those used in McBer's Comprehensive Cognitive Assessment Battery and the Ennis-Weir Critical Thinking Essay Test, can also yield unidimensional measures of the development of critical thinking skills along a continuum. At the same time, objective measures, such as the Erwin Scale of Intellectual Development, can provide multidimensional representations of students' critical thinking abilities.

Although all of the instruments discussed thus far differ in the types of higher-order thinking skills they assess and how those skills are measured, all of these instruments reflect the belief that growth in critical thinking abilities is additive and occurs along a continuum. Measures such as the Reflective Judgment Interview, Measure of Epistemological Reflection, and even Kohlberg's measure of moral development represent a fundamentally different orientation in the assessment of critical thinking skills. The Reflective Judgment Interview, for example, uses expert judgment and a seven-stage model of thinking skills to assess student performance.

As the preceding discussion indicates, a variety of models are available to explain how college affects critical thinking and to measure critical thinking outcomes. My next column will focus on some of the traditional measures of critical thinking, such as the Watson-Glaser Critical Thinking Appraisal and the critical thinking subtest of the Collegiate Assessment of Academic Proficiency. Subsequent columns will examine alternative ways of measuring higher-order thinking skills, including the Reflective Judgment Interview.

Assessment Update
March-April 1996
Volume 8, Number 2

References

Baird, L. L. "Diverse and Subtle Arts: Assessing the Generic Outcomes of Higher Education." In C. Adelman (ed.), *Performance and Judgment: Essays on Principles and Practice in the Assessment of College Student Learning*. Washington, D.C.: Government Printing Office, 1988.

Cuban, L. "Policy and Research Dilemmas in the Teaching of Reasoning: Unplanned Designs." *Review of Educational Research*, 1984, *54*, 655-681.

Pascarella, E. T., and Terenzini, P. T. *How College Affects Students: Findings and Insights from Twenty Years of Research*. San Francisco: Jossey-Bass, 1991.

Thorndike, R. L. "The Central Role of General Ability in Prediction." *Multivariate Behavioral Research*, 1985, 20, 241-254.

Suggested Readings

Halpern, D. F. "A National Assessment of Critical Thinking Skills in Adults: Taking Steps Toward the Goal." In A. Greenwood (ed.), *The National Assessment of College Student Learning: Identification of the Skills to Be Taught, Learned, and Assessed*. Washington, D.C.: Government Printing Office, 1994.

McMillan, J. H. "Enhancing College Students' Critical Thinking: A Review of Studies." *Research in Higher Education*, 1987, *26*, 3-29.

Perkins, D., Jay, E., and Tishman, S. "Assessing Thinking: A Framework for Measuring Critical-Thinking and Problem-Solving Skills at the College Level." In A. Greenwood (ed.), *The National Assessment of College Student Learning: Identification of the Skills to Be Taught, Learned, and Assessed*. Washington, D.C.: Government Printing Office, 1994.

Assessment Update
March-April 1996
Volume 8, Number 2

GARY R. PIKE

The Watson-Glaser Critical Thinking Appraisal

As I noted in my last *Assessment Measures* column, this and the next several columns will focus on various measures of critical thinking. Here I examine the Watson-Glaser Critical Thinking Appraisal (CTA) because it represents an archetype of standardized objective tests of critical thinking ability. Others include the critical thinking subtest of the Collegiate Assessment of Academic Proficiency and the California Critical Thinking Test (see Terenzini and others, 1996). Indeed, researchers have found that correlations between the CTA and other standardized critical thinking tests are quite high (Terenzini, and others, 1996).

The CTA has been used extensively to assess the effectiveness of curricula designed to foster critical thinking.

The CTA is an objective measure of an individual's critical thinking ability that has been used extensively to assess the effectiveness of curricula designed to foster critical thinking (see Crites, 1965; Helmstadter, 1965). The 1964 version of the CTA has two forms, YM and ZM. Both forms of the test, which each take approximately 50 minutes to complete, can be administered in either group or individual settings. In addition to a total score, the CTA provides five subscores: *inference*, the ability to distinguish between true and false inferences made from data; *recognition of assumptions*, the ability to identify unstated assumptions in arguments or assertions; *deduction*, the ability to reason deductively and identify relationships among the elements of a deductive argument; *interpretation*, the ability to weigh evidence and distinguish between generalizations that are warranted by the data and those that are not; and *evaluation of arguments*, the ability to distinguish between arguments that are strong or weak and arguments that are relevant or irrelevant.

Research on the reliability of the CTA has found that the internal consistency of total scores is quite acceptable, ranging from the lower to the upper .80s. Reliability estimates for the CTA subscales are somewhat lower, and users should examine carefully the internal consistency of subscores before using them in assessing program effectiveness. Questions have also been raised about whether the CTA is sufficiently difficult for use with college students (see Crites, 1965; Helmstadter, 1965).

Assessment Update
July-August 1996
Volume 8, Number 4

BEST COPY AVAILABLE

Research on the validity of the CTA has focused on its content, factor structure, and relationship to other measures. Many of the most critical reviews of the CTA have focused on the content of the test. Broadhurst (1970), for example, asked college students to conduct an item analysis of the CTA. On the basis of students' evaluations of the items in the test, he concluded that the CTA is not a valid measure of critical thinking ability. In all fairness, however, the content validity of a test of critical thinking depends on how one defines critical thinking. If one is comfortable with a definition of critical thinking that focuses on the ability to evaluate inferences and assumptions, make deductions, and evaluate argument and interpretation, then the CTA has high content validity. If one has a different definition of critical thinking, then the content validity of the CTA is lower.

Analyses of the structure of the CTA have tended to parallel results I have found for many of the popular tests of general education outcomes. That is, factor analysis results clearly indicate the presence of a dominant general factor, but test items do not load on factors that are consistent with the test developers' a priori scales and subscales (see Follman, Miller, and Burg, 1971). This finding, coupled with relatively low reliability estimates for the subscores, suggests that assessment professionals should exercise caution in their use and interpretation of the CTA subscores.

The content validity of a test of critical thinking depends on how one defines critical thinking.

Several studies have also examined the correlations between CTA scores and other measures in an effort to establish the convergent and discriminant validity of the test. In general, research has found that scores on the CTA are moderately correlated with verbal scores on intelligence tests such as the Weschler Adult Intelligence Scale. These findings led Westbrook and Sellers (1967) to conclude that the CTA measures critical thinking abilities that are not measured by the verbal components of intelligence tests. While some studies have shown that scores on the CTA are higher for students who have participated in specialized critical thinking courses in college (for example, Wilson and Wagner, 1981), other studies have shown that CTA scores are relatively insensitive to college courses that are intended to enhance critical thinking (see Baird, 1988). Based on his review of the literature, Baird cautioned researchers that the CTA may measure a relatively stable construct and as a result the test may be insensitive to educational effects.

Helmstadter's (1965, p. 256) conclusion seems as true today as 30 years ago: "The Watson-Glaser Critical Thinking Appraisal represents a highly professional attempt to measure an important characteristic. And, while there may be some flaws in the test, it is doubtful whether a significantly better measure will be found until there is a major breakthrough either in test technology or in our understanding of the 'thinking' process."

Assessment Update
July-August 1996
Volume 8, Number 4

References

Baird, L. L. "Diverse and Subtle Arts: Assessing the Generic Outcomes of Higher Education." In C. Adelman (ed.), *Performance and Judgment: Essays on Principles and Practice in the Assessment of College Student Learning*. Washington, D.C.: Government Printing Office, 1988.

Broadhurst, N. A. "An Item Analysis of the Watson-Glaser Critical Thinking Appraisal." *Science Education*, 1970, *54*, 127-132.

Crites, J. O. "Test Reviews: Watson-Glaser Critical Thinking Appraisal." *Journal of Counseling Psychology*, 1965, *12*, 328-331.

Follman, J., Miller, W., and Burg, I. "Statistical Analysis of Three Critical Thinking Tests." *Educational and Psychological Measurement*, 1971, *31*, 519-520.

Helmstadter, G. C. "Watson-Glaser Critical Thinking Appraisal." *Journal of Educational Measurement*, 1965, *2*, 254-256.

Terenzini, P. T., Springer, L., Yaeger, P., Pascarella, E. T., and Nora, A. "First-Generation College Students: Characteristics, Experiences, and Cognitive Development." *Research in Higher Education*, 1996, *37*, 1-22.

Westbrook, B. W., and Sellers, J. R. "Critical Thinking, Intelligence, and Vocabulary." *Educational and Psychological Measurement*, 1967, *27*, 443-446.

Wilson, D. G., and Wagner, E. E. "The Watson-Glaser Critical Thinking Appraisal as a Predictor of Performance in a Critical Thinking Course." *Educational and Psychological Measurement*, 1981, *41*, 1319-1322.

The California Critical Thinking Skills Test

Students enrolled in courses designed to enhance critical thinking skills showed significant improvement in their CCTST performance.

Assessment Update
March-April 1997
Volume 9, Number 2

Developed by Peter Facione, the California Critical Thinking Skills Test (CCTST) is one of several recently developed objective measures of students' critical thinking abilities. It grew out of a two-year Delphi research project sponsored by the American Philosophical Association. The expert panel for the Delphi process included 46 individuals who were actively engaged in critical thinking assessment, education, and research. The report on the panel's work focused on generic critical thinking skills appropriate for first-year students and sophomores completing their general education programs.

The original form of CCTST (Form A) consists of 34 items representing five skills identified by the American Philosophical Association's Delphi process: interpretation, analysis, evaluation, inference, and explanation. A sixth critical thinking skill identified by the panel, metacognition, is not measured by the test. Form B of CCTST was developed by rewriting 28 of the 34 items on Form A. Changes in Form B include differences in names, concepts, and contexts; the types of problems and specific critical thinking skills being assessed were not changed.

The 34 items on both forms of CCTST provide a total score and three subscores: analysis, evaluation, and inference, while 30 items can be used to calculate subscores for inductive and deductive reasoning. In his first technical report, Facione (1990a) noted that the KR20 reliability estimates for the total score ranged from 0.68 to 0.69. Jacobs (1995) reported that alpha reliabilities for Forms A and B were 0.56 and 0.59, respectively. Using the Spearman-Brown reliability formula to estimate the internal consistency of subtests composed of 34 items, Jacobs found that reliability estimates ranged from 0.14 for analysis to 0.68 for deductive reasoning on Form A and from 0.42 for analysis to 0.71 for deductive reasoning on Form B.

Jacobs (1995) reported that the item means and patterns of item intercorrelations for the two forms were significantly different. However, these differences may have been the result of differences in test takers. Although

BEST COPY AVAILABLE

students were randomly assigned to test forms, it is not possible to say whether item differences were the product of nonequivalent groups or differences in test forms.

In a series of technical reports, Facione (1990b, 1990c) investigated the convergent and discriminant validity of Form A. In his first study, he found that students enrolled in courses designed to enhance critical thinking skills showed significant improvement in their CCTST performance, while students enrolled in courses not so designed did not. Average total scores for students ranged from 15.4 to 16.1 on the pretest, and from 15.5 to 17.4 on the posttest. The 95% confidence interval for groups showing significant gains on CCTST ranged from 1 to 2 points. Given a standard deviation of approximately 4.5 points, the effect sizes for the lower and upper bounds of significant changes ranged from 0.22 to 0.44 standard deviation units.

In subsequent studies, Facione reported that pre- and posttest scores on CCTST were significantly correlated with college grade point average, Scholastic Achievement Test (SAT) verbal and quantitative scores, Nelson-Denny Reading Test scores and English and mathematics placement test scores. CCTST pretest and posttest scores were both significantly correlated with the number of English courses taken in high school, and pretest scores were also significantly correlated with the number of high school mathematics courses taken. In his research Facione found that CCTST scores were significantly related to reading ability as measured by the Nelson-Denny Reading Test. It is not surprising that native speakers of English scored significantly higher than non-native English speakers. Consistent with Facione's research, Jacobs (1995) found that total scores and subscores on Forms A and B were significantly related to SAT verbal and quantitative scores, with the strongest relationships found between CCTST and SAT verbal scores.

Differences in CCTST performance are also related to instructor characteristics.

Although gender and academic major were not significantly related to students' pretest scores, Facione found that both were related to posttest scores. Specifically, males tended to score higher than females, while students majoring in English, the humanities (excluding the performing arts), mathematics, and engineering scored much higher than students in business and in the performing arts. Students in English, the humanities, mathematics, and engineering also scored higher than students in the natural and social sciences.

Differences in CCTST performance are also related to instructor characteristics. Specifically, number of years of college teaching experience and number of sections of critical thinking courses taught in the prior three years were positively related to CCTST scores. Tenured versus nontenured status, full-time versus part-time employment status, doctoral versus non-doctoral degree, and instructor's gender were not related to CCTST performance. To the extent that teaching effectiveness is related to teaching experience, these findings suggest that performance on the CCTST can be enhanced by the quality of instruction as well as coursework itself.

In sum, CCTST does seem to be sensitive to the educational effects of coursework specifically designed to enhance the critical thinking skills around which the test was designed. However, several questions about the test remain to be answered. These center around the psychometric characteristics of Forms A and B, and the influences of reading ability and general verbal ability on test scores. CCTST users would be well advised to examine carefully the reliabilities and item correlations of total scores and subscores for their student populations. Some scores may not have sufficient internal consistency to provide reliable measures of critical thinking at a given institution.

Users should also examine carefully the relationships between entering verbal ability and CCTST scores in order to determine whether changes in test performance represent college effects of uniform change in a stable construct, such as reading or general verbal ability. Even though instructor effects tend to indicate that gains on CCTST are related to quality of instruction, institutions should collect data on both coursework and instruction to distinguish between the two effects. This is particularly important when the goal of assessment is to evaluate the quality of an educational program.

Several questions about the test remain to be answered.

References

Facione, P. A. *The California Critical Thinking Skills Test-College Level Technical Report 1: Experimental Validation and Content Validity.* Millbrae: California Academic Press, 1990a. (ED 015818)

Facione, P. A. *The California Critical Thinking Skills Test-College Level Technical Report 2: Factors Predictive of Critical Thinking Skills.* Millbrae: California Academic Press, 1990b. (ED 015819)

Facione, P. A. *The California Critical Thinking Skills Test-College Level Technical Report 3: Gender, Ethnicity, Major, Critical Thinking Self-Esteem, and the CCTST.* Millbrae: California Academic Press, 1990c. (ED 015962)

Jacobs, S. S. "Technical Characteristics and Some Correlates of the California Critical Thinking Skills Test Forms A and B." *Research in Higher Education*, 1995, 36, 89-108.

Assessment Update
March-April 1997
Volume 9, Number 2

GARY R. PIKE

The Reflective Judgment Interview

My previous *Assessment Measures* column profiled the Watson-Glaser Critical Thinking Appraisal. Here I describe an instrument from the opposite end of the measurement spectrum, the Reflective Judgment Interview (RJI) (Kitchener and King, 1981). Despite the fact that the RJI is an extremely labor-intensive measure, Wood (1995) has identified more than 30 studies that have made use of the RJI in research on students' intellectual development and critical thinking.

More than 30 studies have made use of the RJI in research on students' intellectual development and critical thinking.

The RJI is based on a model of intellectual development that describes students' abilities to (1) attend to multiple perspectives or problem solutions, (2) produce defensible rationales in support of multiple solutions, and (3) evaluate the various rationales for problem solutions. These three abilities are represented by seven stages of reflective judgment.

At the first stage of reflective judgment, knowledge is assumed to be absolute and concrete, and knowledge is "knowable" based only on direct observation. At the second level, knowledge is still assumed to be absolute and concrete. However, this knowledge can be based on statements of authorities as well as direct observation. At the third level, knowledge is assumed to be absolute, but it may not be immediately available. Consequently, personal beliefs are seen as a valid source of knowledge. At stage four, knowledge is uncertain, and the rationales for the knowledge are based on evidence and reason. Knowledge is uncertain at stage five as well. Moreover, evidence and reason are recognized as context-specific. The quality of evidence in support of knowledge claims is also evaluated as strong-weak, relevant-irrelevant, etc. Individuals at the sixth stage view knowledge as constructed from individuals' perspectives, and interpretations can be evaluated by comparing evidence and opinion from multiple perspectives. The seventh and highest level in the reflective judgment hierarchy views judgments as the outcomes of processes of inquiry. Explanatory value, risks of incorrect interpretations, and consequences of alternative judgments are criteria used to evaluate judgments and conclusions.

Assessment Update
September-October 1996
Volume 8, Number 5

The RJI consists of a series of brief, complex issues and dilemmas that represent ill-structured problems dealing with on-going controversies. Typical issues and dilemmas include evolution versus creation, benefits versus dangers of food additives, and the origin of the pyramids. Students are presented an issue or dilemma and then asked to explain their positions via a series of structured probe questions. These questions are asked by a trained and certified interviewer. Students' responses are transcribed and evaluated by two trained raters using a complex, three-digit scoring system. Details of administration and scoring procedures are presented in greater detail by Kitchener and King (1981, 1985) and more recently by King and Kitchener (1994).

Based on his review of studies using the RJI, Wood (1995) reported that median interrater agreement (within dilemmas) is 77%. Forty percent of the studies reported interrater agreement rates greater than 85%, and one-fourth of the studies reported interrater agreement rates in excess of 90%. King and Kitchener (1994) noted that three of the four studies with interrater agreement rates below 70% used samples of older adult learners and adult nonstudents. King and Kitchener also reported that the median alpha reliability (across dilemmas) of RJI scores is .85 and ranges from .50 to .96.

Students are presented an issue or dilemma and then asked to explain their positions via a series of structured probe questions.

Because the RJI is based on a developmental stage model, it is critical that the seven levels in the RJI scoring scheme represent sequential stages of development. Based on his reanalysis of the actual data from studies using the RJI, Wood (1995, p. 43) concluded that "the results of the sequentiality analyses show that the Reflective Judgment Interview and scoring system documents a complex developmental sequence. Non-sequential responses are quite rare." King and Kitchener (1994) also reported that growth in reflective judgment (that is, movement through stages) is gradual and not due to test-retest effects.

Performance on the RJI appears to be strongly related to the educational attainment of the students being tested. Wood (1995) reported that samples of students in the early stages of their undergraduate careers tend to score about 3.5 on the RJI, while upperclass undergraduates score at level 4. Beginning graduate students tend to score at about 4.5, and advanced doctoral students tend to score at about a 5.5 level.

In his conclusions, Wood (1995) noted that in his research he found that reflective judgment scores tend to be higher for students from small, selective institutions, although these differences may be due to the entering ability of the students (that is, selectivity) rather than institutional size. Wood also noted that patterns of RJI scores differ by level of educational attainment. Specifically, the effect of area of study on RJI scores is more pronounced with graduate student samples than with undergraduates samples.

Results to date strongly indicate that use of the RJI to measure student abilities in relation to ill-structured problems generally produces reliable and valid results. However, some caveats are worth mentioning. First, while the RJI produces highly reliable scores with traditional-age students, questions remain about the reliability of the measure when samples include older adult learners and nonstudent adults. Second, the strong relationship of RJI scores to level of educational attainment raises questions about how sensitive the instrument is to specific educational interventions. Finally, additional research is needed on the effects of expertise and field of study on the variability of student performance on the RJI, particularly at higher levels of educational attainment.

Results to date strongly indicate that use of the RJI to measure student abilities in relation to ill-structured problems generally produces reliable and valid results.

References

King, P. M., and Kitchener, K. S. *Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults*. San Francisco: Jossey-Bass, 1994.

Kitchener, K. S., and King, P. M. "Reflective Judgment: Concepts of Justification and Their Relationship to Age and Education." *Journal of Applied Developmental Psychology*, 1981, *2*, 89-116.

Kitchener, K. S., and King, P. M. *Reflective Judgment Scoring Manual*. 1985.

Wood, P. K. *A Secondary Analysis of Claims Regarding the Reflective Judgment Interview: Internal Consistency, Sequentiality and Intra-Individual Differences in Ill-Structured Problem Solving*. Columbia: Department of Psychology, University of Missouri, 1995.

Critical Thinking Assessment at the University of Missouri, Columbia

There is little agreement on exactly what components should be included in an operationalization of critical thinking, and the relationships among such components remain unclear.

Guest Columnist: Phillip Karl Wood, professor of psychology at the University of Missouri, Columbia.

Many of the calls for educational reform at the postsecondary level focus on the need for college students to reason complexly about issues that have no single correct answer. In addition, mission statements of colleges and universities contain such references when describing the practical benefit of higher education. Often, these calls have focused on cultivating student awareness of difficult real-world problems, evaluating the relative merits of proposed solutions (often within and across different curriculum contexts for understanding the problem at hand), and justifying a chosen position in a rational, defensible manner.

Frequently these broadly defined benefits of higher education are operationalized under the construct of *critical thinking*. Many operationalizations of this construct have appeared in the literature, including such instruments as the Watson-Glaser Critical Thinking Appraisal (Watson and Glaser, 1952, 1994), the Cornell Critical Thinking Test (Ennis and Millman, 1971), the California Critical Thinking Skills Test (Facione and Facione, 1994), and the California Critical Thinking Disposition Inventory (Facione, Facione, and Giancarlo, 1994). Some of these measures have been criticized because they merely present real-world examples of syllogistic reasoning, document the ability of students to articulate implicit assumptions in an inductive argument, over-rely on procedural or topical knowledge of the problem, or confine reasoning about a problem to only one context. In addition, there is little agreement on exactly what components should be included in an operationalization of critical thinking, and the relationships among such components remain unclear (for example, Skinner, 1971).

The Reflective Judgment model has been used to document the increasingly sophisticated justifications of students when they are asked to reason about real-world problems that do not permit single correct answers (also termed *ill-structured* problems by Wood, 1983). These issues, often called *dilemmas*,

Assessment Update
November-December 1997
Volume 9, Number 6

BEST COPY AVAILABLE

differ from the problems considered in many measures of critical thinking in that they involve issues about which reasonable experts could disagree and for which there are answers to only a relative, as opposed to absolute, degree of certainty. One such dilemma is presented in Exhibit 1.

One approach to measuring reflective judgment utilizes the Reflective Judgment Interview (Kitchener and King, 1981). This method involves presentation of four such dilemmas to students within a semi-structured interview format. The semi-structured interview asks the student to express his or her opinions about the issue (either in favor of one position or another or taking no position), but with special emphasis on the justification for the expressed opinion. Additional questions are designed to assess the individual's underlying epistemological assumptions when explaining whether and how one can ever successfully resolve this issue, whether and on what grounds some opinions about the issue can be better than others, as well as whether and how individuals (particularly experts) can disagree. Transcripts of these interviews are then scored by trained, certified raters. Detailed descriptions of the levels of reflective judgment as well as interview and scoring procedures can be found in King and Kitchener (1994) and Wood (1997).

Exhibit 1. Reflective Judgment
Dilemma Topic: Artificial Sweeteners

People often have to make decisions that may affect their health, such as whether to eat foods or drink beverages that contain artificial sweeteners. There have been conflicting reports about the safety of these additives. For example, some studies have indicated that even in small amounts artificial sweeteners (such as Nutrasweet) can cause health problems, making foods containing them unsafe to eat. Other studies, however, have indicated that even in large amounts, artificial sweeteners do not cause health problems, and that the foods containing them are safe to eat.

Use of the Reflective Judgment Interview is impractical for many assessment professionals because of the prohibitive expense associated with interviewing, transcribing, and rating the data.

Toward a More Easily Used Measure of Reflective Judgment

Use of the Reflective Judgment Interview as a measure of epistemic cognition about specific ill-structured problems, although an internally consistent and well calibrated instrument, is impractical for many assessment professionals because of the prohibitive expense associated with interviewing, transcribing, and rating the data. Although efforts have been made recently to develop and score student essays along lines inspired by the Reflective Judgment model (for example, Christen, Angermeyer, Davison, and Anderson, 1997), Patricia King, Karen Kitchener, Cindy Lynch, and I perceived the need for the development of assessment instruments that could be easily administered in a group setting and quickly scored, and that would allow estimation of reasoning about ill-structured problems tied to specific curriculum issues. Although this research is currently underway, it is appropriate to share some of the considerations in the research and instrument design that have guided our efforts. Specifically, I describe here some of the measures we have taken in our efforts to develop and validate our assessment instruments, and, where appropriate, I describe provisional findings of our research to date.

Which Reflective Judgment Skills Should Be Measured?

Because the semi-structured format of the Reflective Judgment Interview requires the participant to produce a reasoned statement of beliefs on the dilemma along with an elaborated view of the epistemic assumptions underlying the position, it did not seem reasonable to ask participants to select among several summary statement responses to a dilemma. Prior research has shown that participants are unable to correctly interpret such summary statements and frequently paraphrase such material in terms of their own epistemic assumptions or reinterpret focus only on one idea or phrase to the exclusion of the general rationale (Kitchener, Lynch, Fischer, and Wood, 1993). Moreover, the ability of individuals to produce justifications for their beliefs is probably quite different from their ability to select and affirm justifications provided to them (commonly referred to as the "production-recognition distinction" in performance; see Rest, 1979). For these reasons, my colleagues and I are currently exploring two approaches to the assessment of epistemic cognitions about ill-structured problems.

The RJET appears to predict a small, but significant, proportion of the variance in course grades in Introductory Psychology above and beyond that explained by ACT composite score or high school core GPA.

In the first measure, termed the Reflective Judgment Essay Test (RJET), students are asked to play the role of a teacher in evaluating pairs of short essay responses to a dilemma topic. Each essay response has been written to correspond to a particular level in the Reflective Judgment model. Students are asked to discriminate between these two prototypical essays by means of Likert-format items for each of six sets of essay comparisons. Specifically, students are asked to rate such things as which essay demonstrates more complex reasoning, which essay is more reasoned and systematic, and which essay reflects a better understanding of how scientists think about the issue at hand. Multiple Likert items are used for each essay comparison in order to assess internal consistency, and multiple comparisons are presented in order to present a variety of Reflective Judgment levels.

Initial analysis of scores based on this instrument has revealed that when first-year undergraduates enrolled in introductory psychology courses are matched to seniors majoring in psychology on academic aptitude (operationalized as American College Test [ACT] composite score) or prior academic achievement (defined as high school core grade point average [GPA]), the RJET scores of psychology seniors are markedly higher than those of the first-year students. No differences have been found on the RJET by sex or race so far. Overall, the RJET appears to predict a small, but significant, proportion of the variance in course grades in Introductory Psychology above and beyond that explained by ACT composite score or high school core GPA. Curiously, however, the RJET appears to predict grade in Introductory Psychology better for men than it does for women. The pattern of these results and their interpretation is the object of ongoing work.

Assessment Update
November-December 1997
Volume 9, Number 6

The second Reflective Judgment instrument, the Reasoning about Current Issues Test (RCIT), presents the student with a short dilemma topic such as that presented above, and asks the student to write a few sentences telling, for example, how it is possible that experts could disagree about such an issue. Following this, students are presented several short summary statements written to correspond to various Reflective Judgment levels. Students indicate the degree to which such summary statements correspond to how they think about the issue, and they are then asked to rate the top three statements that are most like their views on the question. This format is repeated for each of five different dilemma topics. The format of this instrument is somewhat similar to Rest's (1979) Defining Issues Test of moral reasoning, except that students are asked to write summary statements, and the items of the instrument reflect epistemological rather than moral or ethical concerns.

Choices regarding the number and type of summary statements for the RCIT were made in response to careful "think aloud" piloting of the instrument. First, some subjects appeared to endorse items in the instrument based on the vocabulary or evident sophistication of the item rather than the idea being expressed. (This was consonant with Kitchener, Lynch, Fischer, and Wood's [1993] findings regarding interpretation of Reflective Judgment summary statements.) For this reason, each dilemma also contains a summary statement that uses complex vocabulary but is meaningless. Instructions for the instrument inform students that such items exist and that they are to mark these statements as meaningless in the instrument. In practice, this modification had the added benefit that some students classified more advanced-level responses as meaningless, resulting in a clearer discrimination of Reflective Judgment ability. When first-year students were matched to Psychology Department majors on the basis of sex and either ACT composite or high school rank, first-year versus senior differences appeared more pronounced for the RCIT than for the RJET.

It is not enough to know that there are first-year versus senior differences on the proposed measures because numerous alternative explanations could account for the effect.

It is not enough to know that there are first-year versus senior differences on the proposed measures because numerous alternative explanations could account for the effect. Pending further refinement of the instruments and the collection of carefully designed longitudinal data, it seemed reasonable to gather measures that could address such competing explanations. Competing constructs include social desirability (as measured by the Marlowe-Crowne Social Desirability Score; Crowne and Marlowe, 1964), personality correlates of problem solving such as Need for Cognition (Cacioppo, Petty, Feinstein, and Jarvis, 1996), and other student characteristics such as prior academic achievement (as measured by high school core GPA or high school percentile rank) and academic aptitude (as measured by the ACT composite percentile rank). In addition, given the strong correlation of Reflective Judgment scores with educational attainment, first-year students were selected for participation based on number of completed credits as reported by the registrar rather than self-report of educational level.

Finally, given the possible differential effects of academic aptitude and prior academic achievement on subsequent Reflective Judgment ability, it was decided to match subjects in terms of such student characteristics rather than merely specify such variables as covariates. Substantial prior research indicates that gains over time in Reflective Judgment are more pronounced for more able students (as reflected in cross-institution comparisons of Reflective Judgment Interview scores [Wood, 1997] and unpublished analyses of academic aptitude in King and Kitchener's [1994] original ten-year longitudinal study). Although this strategy had the effect of dramatically reducing many of the observed differences between first-year students and seniors, we have much more confidence that the obtained differences that persist under a matched-groups design will persist in subsequent longitudinal replications.

Substantial prior research indicates that gains over time in Reflective Judgment are more pronounced for more able students.

References

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., and Jarvis, W.B.G. "Dispositional Differences in Cognitive Motivation: The Life and Times of Individuals Varying in Need for Cognition." *Psychological Bulletin*, 1996, *119*, 197-253.

Christen, R., Angermeyer, J., Davison, M. L., and Anderson, K. "Performance Assessments of Critical Thinking: The Reflective Judgment Approach." *Research Practice*, 1997, *2*, 1, 3-8.

Crowne, D. P., and Marlowe, P. *The Approval Motive*. New York: Wiley, 1964.

Ennis, R. H., and Millman, J. Cornell *Critical Thinking Test Manual*. Champaign: Critical Thinking Project, University of Illinois, 1971.

Facione, P. A., and Facione, N. C. *The California Critical Thinking Skills Test: Manual Forms A and B*. Millbrae, Calif.: California Academic Press, 1994.

Facione, P. A., Facione, N. C., and Giancarlo, C.A.F. *The California Critical Thinking Disposition Inventory Manual*. Millbrae, Calif.: California Academic Press, 1994.

King, P. M., and Kitchener, K. S. *Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults*. San Francisco: Jossey-Bass, 1994.

Kitchener, K. S., and King, P. M. "Reflective Judgment: Concepts of Justification and Their Relationship to Age and Education." *Journal of Applied Developmental Psychology*, 1981, *2*, 89-116.

Assessment Update
November-December 1997
Volume 9, Number 6

Kitchener, K. S., Lynch, C. L., Fischer, K. W., and Wood, P. K. "Developmental Range of Reflective Judgment: The Effect of Contextual Support and Practice on Developmental Stage." *Developmental Psychology*, 1993, *29*, 893-906.

Rest, J. *Development in Judging Moral Issues*. Minneapolis: University of Minnesota Press, 1979.

Skinner, S. B. "The Myth of Teaching for Critical Thinking." *The Clearinghouse*, 1971, *45*, 572-576.

Watson, G. B., and Glaser, E. M. *Watson-Glaser Critical Thinking Appraisal Manual*. Orlando, Fla.: Harcourt Brace, 1952.

Watson, G. B., and Glaser, E. M. *Watson-Glaser Critical Thinking Appraisal Manual: Form S*. San Antonio, Tex.: Psychological Corporation, 1994.

Wood, P. K. "Inquiring Systems and Problem Structure: Implications for Cognitive Development." *Human Development*, 1983, *26*, 249-265.

Wood, P. K. "A Secondary Analysis of Claims Regarding the Reflective Judgment Interview: Internal Consistency, Sequentiality, and Intra-Individual Differences in Ill-Structured Problem Solving." In J. Smart (ed.), *Higher Education: Handbook of Theory and Research*. Vol. 12. Bronx, N.Y.: Agathon, 1997.

Using Surveys to Measure Critical Thinking Outcomes

A variety of surveys of college outcomes are available commercially.

Assessment Update
May-June 1997
Volume 9, Number 3

In recent columns, I have reviewed a variety of different tests of students' critical thinking abilities. While these measures can provide important information about the effects of college on critical thinking, they tend to be expensive to administer (both in time and money), are strongly related to students' general intellectual abilities, and can be relatively insensitive to the effects of programs designed to improve students' critical thinking skills. Given the limitations of these tests, assessment practitioners may be tempted to consider using surveys to gather information about the development of critical thinking abilities during college. In this column, I examine some of the issues involved in using surveys to measure the development of students' critical thinking abilities.

In two studies, I found that self-reports of students' learning and development during college can be used as proxies for more traditional tests of learning and intellectual development (see Pike, 1995, 1996), with two important caveats. First, there should be a high degree of content overlap between self-reports and objective tests. Second, both self-reports and test scores are strongly influenced by method-specific variance that tends to mask relationships between self-reports and objective achievement tests. Both content overlap and method-specific effects can and should be taken into account when surveys are used to measure the development of critical thinking abilities.

A variety of surveys of college outcomes are available commercially. I am most familiar with the College Outcomes Survey developed by the American College Testing (ACT) program, the College Student Experiences Questionnaire (CSEQ) developed by C. Robert Pace, the alumni and student surveys used by Tennessee in its statewide performance funding program, and the first-year and senior surveys developed at the University of Missouri, Columbia. All of these surveys contain items designed to elicit students' perceptions of their growth and development during college. However, the extent to which they measure the development of critical thinking abilities varies greatly.

BEST COPY AVAILABLE

The ACT College Outcomes Survey, for example, provides very little information about gains in critical thinking. Most questions deal with the acquisition of content knowledge (for example, acquiring a well-rounded general education, and becoming aware of global issues) and basic skills (for example, developing leadership skills and learning to manage finances). While some of the gains questions on the survey could potentially be used to measure critical thinking abilities, Graham and Cockriel (1995) concluded that the outcomes domains represent instead intrapersonal development, personal valuing, social leadership, and civic involvement.

Both the student and alumni surveys developed as part of the Tennessee performance funding program contain items that could potentially be used to measure critical thinking skills (for example, defining and solving problems). Here again, research has shown that the gains questions assess content knowledge and skills rather than critical thinking (see Pike, 1990, 1991). Items related to critical thinking are subsumed within a math and science reasoning domain.

Surveys developed at the University of Missouri, Columbia, also contain several items that were intended to measure gains in students' critical thinking abilities. The questions include skills related to defining and solving problems, thinking analytically and logically, and seeing relationships among ideas. It is noteworthy that these items form a clear critical thinking scale for seniors, but not for first-year students. While extrapolating trends from different samples of first-year students and seniors is dangerous, it is possible that the recognition of how college affects critical thinking develops over the course of a college career. Thus, surveys may be of limited use in identifying the effects of college on the critical thinking abilities of first-year students because these students are not able to identify clearly the gains they have made in critical thinking.

> *Available surveys of student learning focus on the gains in critical thinking that are made during college. They do not assess a student's absolute ability to think critically.*

One of the most widely used measures of students' college outcomes is the CSEQ. The CSEQ contains a variety of questions related to the development of higher-order intellectual skills (such as the ability to think analytically and the ability to put ideas together). Moreover, research has found that these items form a reliable measure of gains in critical thinking that is related to students' college experiences (Pace, 1995). Perhaps the most serious limitation of the CSEQ, and other commercially available surveys, is that the two or three items designed to measure gains in critical thinking provide an extremely narrow view of critical thinking. This narrow perspective may fail to capture many of the ways in which college enhances students' critical thinking abilities.

It is also important to realize that available surveys of student learning focus on the gains in critical thinking that are made during college. They do not assess a student's absolute ability to think critically. While a focus on gains is assumed to remove the effects of entering ability on outcomes, it does not permit clear judgments about what students know and can do. Likewise, a focus on the gains made during a particular year does not allow comparisons of critical thinking attainment over time.

Given the limitations of commercially available surveys, assessment practitioners may be interested in developing their own surveys to measure students' abilities to think critically. If that option is selected, then the first question should be, "What will be measured?" In line with my previous comment, institutions must first decide whether they want to measure gains in critical thinking or absolute levels of critical thinking ability (see Pike, 1996, for examples of how survey questions can be structured to measure levels of absolute attainment). A second question is, "What aspects of critical thinking will be measured?" As noted earlier, the degree of content overlap between surveys and achievement tests is very important. Accordingly, assessment practitioners may want to borrow from the test specifications underlying existing critical thinking measures in order to develop survey questions that can be used as proxies for test scores. The work of both King and Kitchener (1985) and Facione (1990) can provide a starting point in developing survey questions dealing with critical thinking.

Given the limitations of commercially available surveys, assessment practitioners may be interested in developing their own surveys to measure students' abilities to think critically.

The second step in constructing self-report measures of critical thinking is the actual development of survey questions. A variety of sources is available to guide the construction of questionnaires. Consequently, I will not address issues related to survey development here. However, I do believe it is very important that assessment practitioners empirically confirm that their surveys provide reliable and valid measures of critical thinking outcomes. A good survey development process does not always yield a good survey.

The third step in developing a survey of critical thinking is to take advantage of the strength of surveys, identifying why outcomes differ by students. This involves including items that measure students' educational experiences. Recently, I completed a research project designed to evaluate the effects of living-learning communities on the development of students' critical thinking skills. My research suggests that the quality of students' interaction with faculty and peers, as well as their involvement in out-of-class activities, has substantial effects on gains in critical thinking. I found that the most important influence on the development of critical thinking abilities is providing students with structured opportunities to integrate information across courses and across in- and out-of-class activities. While I believe that these findings are significant, they are offered only as suggestions for what types of college experience measures might be included in a survey of critical thinking. The decision about what college experience measures to include in a critical thinking survey must be guided by both theory and the unique goals and characteristics of an institution.

Finally, it is important to remember that strong measurement effects in both surveys and achievement tests may mask the effects of college on the development of critical thinking skills. My own research on living-learning communities suggests that the effects of specific educational interventions designed to promote critical thinking tend to be indirect and subtle. For example, learning communities seem to have the greatest

Assessment Update
May-June 1997
Volume 9, Number 3

effect on faculty-student and peer interaction, which enhances opportunities for students to synthesize and integrate information across courses and out-of-class experiences. This, in turn, influences the development of critical thinking abilities. Consequently, the direct effects of learning communities are very small, although the indirect effects are quite substantial.

If my findings regarding the effects of living-learning communities on gains in critical thinking are generalizable to other types of educational interventions, then the insensitivity of existing critical thinking measures may not be a psychometric or measurement problem at all. Instead, modest gains on most measures of critical thinking may simply reflect Pascarella and Terenzini's (1991) conclusion that college effects on learning and development tend to be cumulative and gradual. Unfortunately, claiming that college effects are cumulative and gradual may not be a sufficient response to political problems created by constituents' demands for accountability.

College effects on learning and development tend to be cumulative and gradual.

References

Facione, P. A. *The California Critical Thinking Skills Test-College Level: Experimental Validation and Content Validity.* Report No. 1. Millbrae: California Academic Press, 1990.

Graham, S., and Cockriel, I. "Indexes to Assess Social and Personal Development and the Impact of College." Paper presented at the annual meeting of the Association for the Study of Higher Education, Orlando, Florida, Nov. 1995.

King, P. M., and Kitchener, K. S. "Reflective Judgment Theory and Research: Insights into the Process of Knowing in the College Years." Paper presented at the annual meeting of the American College Personnel Association, Boston, Mar. 1985.

Pace, C. R. "From Good Processes to Good Products: Relating Good Practices in Undergraduate Education to Student Achievement." Paper presented at the annual forum of the Association for Institutional Research, Boston, May 1995.

Pascarella, E. T., and Terenzini, P. T. *How College Affects Students: Findings and Insights From Twenty Years of Research.* San Francisco: Jossey-Bass, 1991.

Pike, G. R. "Dimensions of Alumni Perceptions of Cognitive and Affective Growth During College." Paper presented at the annual meeting of the American Educational Research Association, Boston, Apr. 1990.

Pike, G. R. "Dimensions of Student Academic Development During College: Using Alumni Reports to Evaluate Education Programs." Paper presented at the annual meeting of the Association for the Study of Higher Education, Boston, Oct. 1991.

Pike, G. R. "The Relationship Between Self Reports of College Experiences and Achievement Test Scores." *Research in Higher Education*, 1995, 36, 1-22.

Pike, G. R. "Limitations of Using Students' Self Reports of Academic Development as Proxies for Traditional Achievement Measures." *Research in Higher Education*, 1996, 37, 89-114.

Assessment Update
May-June 1997
Volume 9, Number 3

GARY R. PIKE

College Student Experiences Questionnaire

In the fall of 1989, the Center for Assessment Research and Development at the University of Tennessee, Knoxville, began using the *College Student Experiences Questionnaire* (CSEQ), developed by C. Robert Pace, in a study of factors influencing learning in general education. My experience with this questionnaire and my knowledge of its widespread use in educational research convinced me to devote a column to a profile of the instrument. I will provide a brief overview of the CSEQ, including its rationale and structure and evidence of its reliability and validity.

In *Measuring the Quality of College Student Experiences*, Pace (1988) argues that while outcomes are important, the assessment of educational programs also needs to examine the process of education. Accordingly, he designed the CSEQ to provide measures of the educational process, as experienced by the student.

Much of the literature on program assessment makes the tacit assumption that if student outcomes are not satisfactory the institution is at fault. While this may be true, it is also true that students influence the quality of the education they receive. As Pace points out, what students get out of college depends on what they put into it. Thus, the CSEQ measures the quality of students' effort, as well as the quality of the college environment.

Pace designed the CSEQ to evaluate the whole college environment. The content of the CSEQ covers activities that occur on most college campuses and have been shown to foster learning and development (use of the library, interaction with faculty, membership in clubs and other campus organizations).

The CSEQ contains six parts. The first part consists of questions designed to gather background information, such as age, gender, parents' education, campus residence, major field, and other variables. The second part of the questionnaire focuses on 14 college activities in which the student may have participated. Responses are used to construct a series of Guttman type

Because students' learning and development can occur in or out of the classroom, Pace designed the CSEQ to evaluate the whole college environment.

Assessment Update
Spring 1990
Volume 2, Number 1

scales representing the quality of effort the student has invested in each activity. For example, the 10 questions on use of the library form a scale ranging from "routine, moderately exploratory use" to "increased amount of independent exploration and focused activity" (Pace, 1988, p. 12). Sections three through five are relatively short and ask students to report the amount of reading and writing they have done during the last year, indicate their satisfaction with their college experiences, and evaluate the college environment. The final section of the questionnaire contains 21 questions concerning gains that students have made in college. Factor analysis of these questions indicates that they can be grouped according to five dimensions: personal and social development; science and technology; general education, literature, and the arts; intellectual skills; and vocation.

Research on the CSEQ has shown it to be a highly reliable and valid instrument. For example, studies of the internal consistency of the questions comprising the quality-of-effort scales indicate the reliability ranges from .79 for use of the library and course learning activities to .90 for membership in clubs and organizations and the use of science laboratories.

Research on the CSEQ has shown it to be a highly reliable and valid instrument.

Studies of the validity of the CSEQ show that the quality-of-effort scales are the best predictors of perceived gains from general education, and that these scales are an important predictor of students' cumulative grade-point averages. In addition, the questions on perceived gains and on quality of the college environment are strongly related to students' levels of satisfaction with their college experiences.

Because many institutions have used the CSEQ for assessment, accountability, accreditation self-study, and similar purposes of institutional research and evaluation, a substantial norm group is available for comparisons. A revised third edition of the CSEQ was published in January 1990. Previously published normative results remain appropriate for more than 90% of the contents of the revised third edition. Pace encourages institutions to replicate his research on the validity of the CSEQ to identify relationships between students' effort, the college environment, and educational outcomes. All of this information can be very useful for institutions interested in evaluating the quality of their educational programs.

Note: Regarding the Fall 1989 (vol. 1, no. 3) *Assessment Measures* column, comparing three tests of general education, a reader has noted that both the Academic Profile and the Collegiate Assessment of Academic Proficiency (CAAP) have been modified since the Washington study. This qualification is present in both the *General Report* and the *Technical Report*, but was not included in the executive summary, which was part of the *General Report* and was printed verbatim in the *Assessment Measures* column.

Reference works by C. Robert Pace: *Measuring the Quality of College Student Experiences*. Los Angeles: Higher Education Research Institute, 1988.

The Undergraduates. Los Angeles: CSE, 1990.

Assessment Update
Spring 1990
Volume 2, Number 1

GARY R. PIKE

The Community College Student Experiences Questionnaire

My own bias is that any comprehensive assessment program should focus on what students do (that is, educational processes) as well as on what students learn (that is, educational outcomes). A similar philosophy is reflected in the National Center for Higher Education Management Systems report on using indicators of good practice in undergraduate assessment (see Ewell, Lovell, Dressler, and Jones, 1994). It also is reflected in the research of Pascarella, Terenzini, and their colleagues in connection with the National Study of Student Learning and in many of the surveys I have reviewed for *Assessment Update*. As Pascarella and Terenzini (1991, p. 610) noted in *How College Affects Students*, "One of the most inescapable and unequivocal conclusions we can make is that the impact of college is largely determined by the individual's quality of effort and level of involvement in both academic and nonacademic activities."

The impact of college is largely determined by the individual's quality of effort and level of involvement in both academic and nonacademic activities.

In an earlier issue of *Assessment Update* (Vol. 2, No. 1), I reviewed the College Student Experiences Questionnaire (CSEQ), developed by C. Robert Pace. In that column, I noted that the CSEQ has been used by a variety of four-year colleges and universities and that the quality-of-effort scales in the CSEQ provide important information about what students do in college.

While the CSEQ is a useful assessment tool for four-year institutions, it is not appropriate for two-year institutions. Many of the experiences in the CSEQ quality-of-effort scales (for example, Student Unions and Residence Halls) are not appropriate for two-year institutions. At the same time, experiences that are appropriate for students at two-year institutions (for example, experiences related to the development of vocational skills) are not included in the CSEQ.

Recognizing the limitations of the CSEQ for two-year institutions, Pace, Friedlander, and Lehman developed and published the Community College Student Experiences Questionnaire (CCSEQ) in 1990. Since its publication, the CCSEQ has been administered to more than 18,000 students at 56 community colleges nationwide. From 1990 to 1994, the CCSEQ was

Assessment Update
January-February 1996
Volume 8, Number 1

BEST COPY AVAILABLE

disseminated by Pace at the University of California-Los Angeles. Since 1994, the base of operations for the CCSEQ has been moved to the Center for the Study of Higher Education at the University of Memphis under the directorship of Patricia H. Murrell.

Friedlander, Murrell, and MacDougall (1993) surveyed representatives from 26 community colleges that have administered the CCSEQ. Uses of the CCSEQ have ranged from local institutional assessment and evaluation to statewide reporting and basic research on community college students. Examples of the ways in which CCSEQ data have been used in assessment programs range from making changes in specific courses to establishing new programs institution-wide. Specifically, English Department faculty at one institution found that a large number of students were not utilizing library resources effectively. The faculty subsequently modified several courses to include a component on using the library. Likewise, another school found that student involvement was extremely poor. The institution subsequently developed a component in its new student orientation program that focused on student involvement.

The CCSEQ quality-of-effort scales can be used in accommodating the learning needs of particular subgroups of students.

By far, the most significant change in the CCSEQ that distinguishes it from the CSEQ is the reduction in the number of quality-of-effort scales, from 14 in the CSEQ to 7 in the CCSEQ. The quality-of-effort scales in the CCSEQ are Course Learning; Library Usage; Interaction with Faculty; Student Acquaintances; Art, Music, and Theater; Science Activities; and Vocational Skills.

Like the 14 scales in the CSEQ, the quality-of-effort measures in the CCSEQ consist of Guttman-like items designed to measure both the breadth and depth of student involvement. For example, the quality-of-effort scale on course learning includes questions related to the number of times the respondent participated in class discussions, summarized major points and information from readings or notes, studied course materials with other students, compared different points of view presented in a course, and considered the accuracy and credibility of information from different sources. Measures of internal consistency (alpha reliability) exceed 0.80 for the CCSEQ quality-of-effort scales.

Research on the construct validity of the CCSEQ quality-of-effort scales by Lehman (1992) found that the empirical structure of those scales matched their theoretical structure. More recently, Ethington (1996) found that the empirical structure of the quality-of effort scales was similar, irrespective of the subgroup (males versus females or minority versus majority students) being studied. Based on her findings, Ethington concluded, "Thus, the CCSEQ quality-of-effort scales can be used to provide valid and insightful information for faculty and administrators to use in helping students adapt to the institutional environment and in accommodating the learning needs of particular subgroups of students." In sum, the CCSEQ can be an important component in the assessment programs of two-year institutions interested in evaluating and improving the quantity and quality of student involvement.

Assessment Update
January-February 1996
Volume 8, Number 1

References

Ethington, C. A., and Polizzi, T. B. "An Assessment of the Construct Validity of the CCSEQ Quality-of-Effort Scales." *Research in Higher Education*, 1996, 37, 7.11-7.30.

Ewell, P. T., Lovell, C. D., Dressler, P., and Jones, D. P. *A Preliminary Study of the Feasibility and Utility for National Policy of Instructional "Good Practice" Indicators in Undergraduate Education*. NCES 94-437. Washington, D.C.: Government Printing Office, 1994.

Friedlander, J., Murrell, P. H., and MacDougall, P. R. "The Community College Student Experiences Questionnaire." In T. W. Banta and Associates, *Making a Difference: Outcomes of a Decade of Assessment in Higher Education*. San Francisco: Jossey-Bass, 1993.

Lehman, P. W. CCSEQ: *Test Manual and Comparative Data*. Los Angeles: Center for the Study of Evaluation, University of California, 1992.

Pascarella, E. T., and Terenzini, P. T. *How College Affects Students: Findings and Insights from Twenty Years of Research*. San Francisco: Jossey-Bass, 1991.

ACT
Evaluation/Survey Service

A standard feature of all of the ESS surveys is an additional-item section in which institutions may ask up to 30 locally developed survey items.

Drawing on extensive experience in developing customized survey instruments for colleges and universities, the American College Testing Program (ACT) introduced the ACT Evaluation/Survey Service (ESS) in 1979. Since that time, almost 3 million ESS instruments have been administered at hundreds of institutions in the United States and Canada. According to ACT (1989, p. 2), "the primary purpose of the ESS is to assist educational institutions and agencies in the collection, interpretation, and use of student survey data for such purposes as institutional planning, research, evaluation, and self-study."

At the present time, the ESS makes available 12 different surveys to two- and four-year colleges and universities. These instruments include needs assessment surveys for college students and adult learners, satisfaction surveys for enrolled students, alumni surveys, and surveys of withdrawing and nonreturning students. Each instrument contains multiple parts. For example, the ESS Alumni Survey includes questions about continuing education, college experiences, and employment history, while the Student Opinion Survey includes questions about the quality of college services and the college environment.

A standard feature of all of the ESS surveys is an additional-item section in which institutions may ask up to 30 locally developed survey items. ACT has compiled an "Optional Item Catalog" that institutions may use in selecting and/or developing local items. The only restrictions placed on the types of additional items included on a given questionnaire are that they have 12 or fewer response options and that students may select only one response option for a question. Thus, questions that state "check all that apply" are not permitted.

Assessment Update
July-August 1991
Volume 3, Number 4

Two other standard features of the ESS surveys are the scoring and reporting service and the subgroup option. ACT scans student answer sheets and prepares institutional summary reports on a monthly basis. User norms for public and private colleges are furnished with institutional reports. In addition to institutionwide data, up to 15 subgroups may be designated by an institution and reports will be developed for each subgroup. Features offered by the ESS include reports with descriptions of user norms, computer tapes containing the raw data for an institution, and composite and bound reports. The ESS also provides assistance to institutions in designing evaluation studies, administering surveys, and interpreting, reporting, and using results.

Studies of the reliability of three of the instruments, the Student Opinion Survey, the Entering Student Survey, and the Adult Learner Needs Assessment Survey, indicate that the ESS measures are highly reliable. For example, correlations of responses over two administrations of the Student Opinion Survey have been reported to be .92 for items related to satisfaction with college programs and services and .95 for satisfaction with various aspects of the college environment. Reliability estimates based on institutional (or program) means have ranged from .90 to .98 for the Student Opinion Survey and from .98 to .99 for the Adult Learner Needs Assessment Survey. Group means reliability estimates for the Entering Student Survey have been reported to be .99.

Studies of the reliability of three of the instruments indicate that the ESS measures are highly reliable.

For additional information about the ACT Evaluation/Survey Service, please contact the American College Testing Program, Institutional Services Area, P.O. Box 168, Iowa City, IA 52243.

Reference

American College Testing Program. *ESS: The ACT Evaluation/Survey Service*, Iowa City, Iowa. American College Testing Program, 1989.

ACT ASSET Program

Immediate scoring is possible, which allows the ACT ASSET to be used for walk-in orientation, advising, and registration.

In previous columns, I have focused either on students' self-reports (surveys) or objective measures (tests). The ACT ASSET program is a combination of both methods—a locally administered and scored advising, placement, and retention system. Although the ACT ASSET is advertised as having been designed for two-year institutions, much of the information it provides could be used by four-year colleges and universities as well.

The most efficient setting for administration of the ACT ASSET is a group, but it can also be given to individuals. Immediate scoring is possible, which allows the ACT ASSET to be used for walk-in orientation, advising, and registration.

The ACT ASSET's educational planning form provides self-report data on students' backgrounds (age, gender, ethnic origin), high school performance (coursework and grades), and postsecondary educational experiences (credit hours and degrees earned). The planning form also asks students about their reasons for attending the institution, their educational and career goals, and their special needs (handicaps, basic skills deficiencies, financial need). Several locally developed items may be added to the planning form.

The second part of the ACT ASSET is its testing program. Four types of tests may be included: on basic skills (numerical, reading, language, and study skills), advanced skills (three levels of algebra and advanced language skills), career skills (mechanical reasoning, clerical speed and accuracy, and space reasoning), and locally developed skill measures (up to five scores can be accommodated). Both the advanced and career skills tests can be tailored to individual students. For example, students do not have to take all three advanced algebra tests, and the career skills test should be reserved for students interested in particular occupational or technical programs.

A microcomputer scoring and reporting system is optional for the ACT ASSET. The software for the system provides immediate scoring and reporting of assessment results. In addition, database software is included.

Assessment Update
Fall 1990
Volume 2, Number 3

BEST COPY AVAILABLE

so that institutions can produce mailing lists, mailing labels, and letters for follow-up contacts and retention research. Data can be transferred from the microcomputer to a mainframe computer for further analysis or for archiving.

Comprehensive research reports are available from the American College Testing Program. These include an entering student descriptive report and a returning student retention report. Subgroup reporting is available, so that institutions can target and track specific groups of students. Several norm groups are available, so that campuswide results can be compared to district, state, and national results.

A key element of the ACT ASSET program is the use of "success seminars." These seminars (and the ACT ASSET battery) are based on the assumption that learning and retention are influenced by five factors: entry skills, learning resources, students' awareness of campus services, students' interest, and students' effort.

Data concerning each of the five elements are provided by the ACT ASSET and are used in the success seminars. For example, students' interest is measured by those parts of the educational planning form that are related to educational and career goals. These data are used in the success seminars to complete a faculty-advising recommendations form that helps guide students' curricular choices.

The success seminars also provide a vehicle for further data collection. For example, some information about students' effort is available in the study skills test, but this information could be supplemented in the success seminars by questions about time management.

The materials I have received on the ACT ASSET do not contain information about the psychometric characteristics of the measures. Interested individuals can write to the American College Testing Program (P.O. Box 168, Iowa City, IA 52243) for information about the reliability and predictive validity of the measures. If the tests of basic and advanced skills will be used in making placement decisions, information about setting the levels for passing scores should also be requested.

Institutions can target and track specific groups of students.

Student Reactions to College (SRC) Questionnaire

The SRC survey represents a tradeoff between what students want to say and what administrators and faculty want to hear.

This column is the second in a series on self-report measures of affective college outcomes. It focuses on the Student Reactions to College (SRC) questionnaire developed by the Educational Testing Service (ETS). The SRC is a 150-item survey with versions for both two- and four-year institutions.

The SRC was developed to give students a vehicle for expressing their opinions about their educational experiences to administrators and faculty. At the same time, it gives administrators and faculty an opportunity to hear what students have to say. Given these two goals, it is inevitable that the SRC survey represents a tradeoff between what students want to say and what administrators and faculty want to hear.

The test developers identified a variety of considerations that guided the construction of the SRC, four of which I believe are extremely important. First, the questions in the SRC survey focus on issues that concern students and about which administrators and faculty can take action. As a result, the SRC is a "very applied" questionnaire.

A second consideration in designing the SRC was that it yield group (not individual) scores. Thus, the SRC is designed for program evaluation, rather than student evaluation, and the confidentiality of student responses is assured.

The third factor considered in the design of the SRC was the need to present results in a manner that can be understood by administrators and faculty who are familiar with their academic programs, but who may have little background in either measurement or statistics. Consequently, responses are reported as percentages. Percentage distributions for combined data from several institutions also are presented.

Finally, it was assumed that each question on the SRC should yield unique information. As a result, scales were not developed for the questionnaire.

Assessment Update
Summer 1990
Volume 2, Number 2

although the questions are grouped into content areas (based on the results of a series of factor analyses described in the questionnaire's *Manual for Users*).

The questions in the SRC are divided among four broad categories and 19 specific areas. These categories and areas are: (1) instruction and study habits (quality of instruction, form of instruction, student-centered instruction, academic performance, grading, studying, instructor accessibility, and involvement with faculty); (2) student goals (counseling/advising and planning); (3) student services (programming, registration and scheduling, library/bookstore, rules and regulations, and administrative procedures); and (4) student activities outside class (campus climate, organized student activities, help with living problems, and financial and related problems).

The SRC reports provide summaries of students' responses grouped by type of question and by up to five student subgroups. Subgroups should consist of a minimum of 50 students and that conclusions about an institution be based on a sample of at least 300 students (preferably 400 to 500 students).

Results from the SRC can be used in formal self-study procedures for accreditation or in more informal (program planning and evaluation) efforts. When the results of the SRC are used for accreditation, institutions should first relate each of the 19 content areas covered by the survey to the institution's mission, goals, and objectives. Once this has been accomplished, results from the survey can provide one type of information about the institution's effectiveness in attaining its goals and objectives.

Results from the SRC can be used to identify the strengths and weaknesses of various programs.

In less formal evaluation and planning efforts, results from the SRC can be used to identify the strengths and weaknesses of various programs. More specifically, SRC results can be used to identify programs that are performing poorly or exceptionally well. The results can also be used in planning to guide (and provide empirical support for) decisionmaking. And the SRC results can provide evaluation information about the effectiveness of program changes over time.

As previously noted, the SRC can provide information about student groups, as well as programs. Responses can help identify the extent of student dissatisfaction, and they can be used to identify groups of students who may need special attention. (This is particularly true for the sections of the questionnaire concerned with living problems and financial problems.)

Eleven pages of the *Manual for Users* are devoted to a nontechnical discussion of the reliability and validity of the SRC. Studies at ETS indicate that the reliabilities for all but one content area are in excess of .90. The discussion of validity in the user's manual focuses on the results of the factor analyses used to group questions into content areas. The section on reliability and validity also contains a very useful discussion of sampling procedures and the effect of differences in sample sizes on the dependability of results.

Readers interested in additional information about the Student Reactions to College questionnaire should call or write: ETS College and University Programs, Princeton, NJ 08541-0001. Tel: (609) 734-1162.

Reference

Educational Testing Service. *Student Reactions to College: Manual for Users.* Princeton, NJ: Educational Testing Service, 1978.

Assessment Update
Summer 1990
Volume 2, Number 2

GARY R. PIKE

The Noel-Levitz Student Satisfaction Inventory

Noel-Levitz has earned a national reputation for its National Centers for Enrollment Management and Student Retention. Recently, I became acquainted with the Noel-Levitz *Student Satisfaction Inventory (SSI)*, four-year college and university version. The *SSI* is a 79-item survey designed to measure how satisfied students are and what is important to them.

The items on the *SSI* measure satisfaction and importance in 12 areas: *academic advising effectiveness*, covering the approachability, knowledge, competence, and personal concern for student success of academic advisers; *campus climate*, assessing the extent to which the institution communicates with students effectively and provides experiences that promote campus pride and a sense of inclusion and involvement; *campus life*, measuring students' perceptions of campus policies and the effectiveness of student life programs such as athletics and residence life; *campus support services*, covering the quality of education support programs such as tutoring, libraries, computer labs, and academic and career services; *concern for the individual*, assessing the commitment of administrators, faculty, and staff to treating each student as an individual; *instructional effectiveness*, assessing students' perceptions of the curriculum, teaching effectiveness, and availability of faculty outside class; *recruitment and financial aid effectiveness*, measuring such factors as the knowledge of admissions counselors and the availability and effectiveness of financial aid; *registration effectiveness*, covering various issues associated with the quality and effectiveness of registration and billing; *responsiveness to diverse populations*, including responsiveness to racial and ethnic minorities, students with disabilities, commuters, and nontraditional students; *safety and security*, measuring students' perceptions of their personal safety on campus and the quality of security personnel; *service excellence*, assessing students' perceptions of the attitudes of frontline staff toward students; and *student centeredness*, assessing campus efforts to convey to students that they are important to the institution and the extent to which students feel valued and welcome.

The SSI is a 79-item survey designed to measure how satisfied students are and what is important to them.

Assessment Update
January-February 1995
Volume 7, Number 1

BEST COPY AVAILABLE

According to Noel-Levitz, the *SSI* takes approximately 15-20 minutes for students to complete. In addition, the authors of the survey provide detailed guidelines for picking an appropriate sample size. For example, Noel-Levitz recommends that institutions with 1,000-2,000 students survey 50% of their student body or 600 students, whichever is higher. Institutions with more than 10,000 students are advised to survey 20% of the population or a minimum of 2,500 students, whichever is higher.

Users of the *SSI* receive a campus report including a description of the survey, a summary of results indicating what is most important/satisfying and least important/satisfying to students. Campus data and user norms for the 12 scales and a detailed item analysis are also provided. Users may request two types of target-group reports. The first, the *Comparative Summary Analysis*, provides mean satisfaction and importance scores for scales and items for a predefined subpopulation. The second, the *Single-Group Analysis*, provides in-depth data paralleling the campus report on a particular subgroup. In the order form, Noel-Levitz lists over 60 special populations and provides space for additional campus-defined groups that can be identified in the target-group reports.

D*ata from the* SSI *can be used for campus improvement.*

Data on the reliability and validity of the *SSI* scales are included in the interpretive guide accompanying an institution's campus report. In addition, the interpretive guide includes a discussion of how the data from the *SSI* can be used for campus improvement.

Prices for the *SSI* vary depending on the quantities ordered. Prices range from $1.95 per survey (for orders of 100 to 1,000) to $1.50 per survey (for orders of 2,500 or more). There is also a $50 processing fee, and the cost of a data diskette is $50. Costs of target-group reports are $15 for each *Comparative Summary Analysis* and $25 for each *Single-Group Analysis*.

For more information about the Noel-Levitz *Student Satisfaction Inventory*, please contact Julie Hanschman, Manager SSI, Noel-Levitz, 2101 ACT Circle, Iowa City, IA 52245-9581. Tel.: (800) 876-1117 or (319) 337-4700, Fax: (319) 337-5274.

Assessment Update
January-February 1995
Volume 7, Number 1

Student Goals Exploration

The *Student Goals Exploration* (Version C), developed by the National Center for Research to Improve Postsecondary Teaching and Learning at the University of Michigan, is designed to assess the goals that "students bring to specific introductory courses early in a semester" (Stark and Lowther, 1990). Two other versions of this instrument, which assess students' goals near the end of a term (Version D) and their goals for their major field of study (Version M), are also available.

All three versions of the *Student Goals Exploration* ask questions concerning students' background characteristics, their goals in taking a particular course (or selecting a major), their expectations and study skills in the course (or major), and their goals in attending college. Each set of questions yields several scales that are moderately correlated. Questions regarding students' goals in attending college yield three scales: General Education, Career/Vocation, and Nondirected (Shaw, Stark, Lowther, and Sossen, 1990). Ten items are in the General Education scale, including "To become a more cultured person" and "To learn about interesting things." Five items comprise the Career/Vocation scale, including "To get a better job after college" and "To be able to make more money." Four items make up the Nondirected scale, including "To get away from home" and "To meet family expectations." Estimates of alpha reliability are .83 for General Education, .71 for Career/Vocation, and .69 for Nondirected (Shaw, Stark, Lowther, and Sossen, 1990).

The heart of the survey is the section on students' goals for a particular course or major. Versions C and M yield nine parallel scales: Increase My World Understanding; Develop Personal Independence, Self Understanding; Develop Scientific Investigative Skills; Build Career Success; Develop Reasoning and Decision Skills; Develop Aesthetic, Creative, Leisure Interests; Acquire Useful Expertise; Develop Social Relations Skills; and Develop Basic Study and Information Skills. Alpha reliability estimates for these scales range from .69 to .92 (Stark and Lowther, 1990).

The Student Goals Exploration ***is designed to assess the goals that "students bring to specific introductory courses early in a semester."***

Assessment Update
Winter 1990
Volume 2, Number 4

BEST COPY AVAILABLE

The items concerning students' expectations and study skills provide similar but not precisely equivalent scales for versions C and M. (This absence of equivalent scales is most interesting, since the same items are included in both versions of the survey.) Two parallel scales are available for versions C and M. These are the Confident Scholar and Anxious Diligent Student scales. Although their titles suggest that these scales are opposites, a study using these scales, one that I reviewed recently, indicates that they are not correlated. This finding may be explained by the fact that the Confident Scholar scale measures students' expectations for learning, while the Anxious Diligent Student scale measures students' expectations for performance on tests.

Several studies have been conducted to validate the *Student Goals Exploration*. One study, designed to assess the relationships among students' background characteristics, college-level goals, and course-level goals, found that background characteristics, to some extent, do predict the goals that students bring to college, and the direction of these effects is consistent with other research (Shaw, Stark, Lowther, and Sossen, 1990). The same study found that the three college-level goals are more strongly related to course-level goals than to students' background characteristics. The authors conclude that these results generally support the validity of the *Student Goals Exploration*.

Background characteristics, to some extent, do predict the goals that students bring to college.

In another study, Stark and Lowther (1990) found that the *Student Goals Exploration* discriminates between business majors and liberal arts majors and between students in business courses and those in liberal arts courses. Business students have strong career-oriented goals, while liberal arts students have particularly strong commitments to general education goals. Because similar results were obtained for versions C and M, the authors conclude that this study provides evidence of both the convergent and discriminant validity of the survey. This study also finds statistically significant differences between business and liberal arts students, but those differences are not great enough to justify a stereotype of business students as having narrow educational interests.

For additional information, write to the National Center for Research to Improve Postsecondary Teaching and Learning, 2400 School of Education Building, University of Michigan, Ann Arbor, MI 48109-1259.

References

Shaw, K. M., Stark, J. S., Lowther, M. A., and Sossen, P. L. "Predicting Students' College and Course Goals from Their Background Characteristics." Paper presented at the annual Forum for the Association for Institutional Research, Louisville, KY, May 1990.

Stark, J. S., and Lowther, M. A. "College and Course Goals: Do Business Students Differ from Other Students?" Paper presented at the annual Forum for the Association for Institutional Research, Louisville, KY May 1990.

Assessment Update
Winter 1990
Volume 2, Number 4

GARY R. PIKE

The Consortium on Financing Higher Education's Alumni Survey

Many of the columns I have written for *Assessment Measures* have focused on commercially available outcomes measures. In this column I want to describe an institutionally developed measure, the alumni survey created by the Consortium on Financing Higher Education (COFHE). During the 1992 annual forum of the Association for Institutional Research (AIR), I attended a session describing this instrument and was sufficiently impressed to share the description that I learned.

The COFHE alumni survey contains five parts. The first asks graduates to report on their education and employment experiences since graduation. Questions include whether alumni are seeking graduate degrees, if they are employed, the occupations in which they are employed, and factors influencing their career choices.

The second part of the survey elicits an evaluation of undergraduate experiences. Questions ask about involvement in campus activities during college and the contributions of involvement to personal development. Several questions about experiences in the major are included, as is a series of questions exploring the gains/progress made during college. Alumni are asked to indicate the importance of gains in their current life and to rate the extent to which college experiences contributed to these gains. The second part of the survey concludes with items about general satisfaction with college.

The third part of the survey asks graduates to report personal and financial data, while the fourth part deals with how they financed their college education. Specific questions in the fourth section identify sources of financial assistance and the effects that scholarships, grants, and loans may have had on current interests and activities. This type of information is seldom asked in alumni surveys, but could prove useful to institutions interested in examining the effects of college financial aid programs. The final section of the survey focuses on alumni participation in community and cultural activities since graduation.

Questions in the fourth section identify sources of financial assistance and the effects that scholarships, grants, and loans may have had on current interests and activities.

Assessment Update
September-October 1992
Volume 4, Number 5

BEST COPY AVAILABLE

At the 1992 AIR forum, Larry Litten and Joseph Pettit presented the results of a preliminary study using the COFHE survey. The survey was mailed to 13,241 alumni at nineteen private, selective institutions. Two mailings of the questionnaire produced a response rate of 48 percent, ranging from 38 to 58 percent across institutions. Results for 5,279 individuals from thirteen institutions were presented. The institutions included seven universities (three Ivy League institutions), three coed colleges, and three women's colleges.

Factor analysis of the items concerning the importance of gains made during college identified six dimensions: (1) intellectual development (the ability to put ideas together, analytical/logical thinking, and independent learning); (2) development of values (understanding self, understanding and getting along with different kinds of people, developing ethics and values); (3) development of leadership (ability to lead groups, speak before groups, and function as a team member); (4) humanistic development (knowledge of literature, understanding of music/art/drama, knowledge about other parts of the world, and awareness of different philosophies and cultures); (5) development of practical skills (skills and knowledge for a job, information relevant to a career, and background for further education); and (6) development of scientific knowledge (understanding of new scientific and technical developments, understanding of scientific experimentation, awareness of the consequences of new science applications, and quantitative thought).

The survey was mailed to 13,241 alumni at nineteen private, selective institutions.

Perhaps the most interesting finding of the preliminary study is the relationships among key items on the survey. Using path analysis, researchers found that involvement in university activities was positively related to perceptions of gains made during college. Those gains in turn were related to satisfaction with college. Both perceived gains during college and general satisfaction with college were related to participation in civic and cultural activities. These relationships and others are discussed in greater detail in material from the presentation.

While I believe that the COFHE alumni survey is a promising assessment instrument, I would not recommend that an institution simply use it without adapting it for its own programs, as the survey was developed to provide information about issues of interest to private, selective institutions in the Northeast. However, I do believe that many of the items on the survey may be useful to other institutions, and I would strongly encourage readers interested in developing (or revising) their own alumni surveys to obtain a copy of the COFHE survey and supporting materials. I believe that this survey may provide assessment practitioners with some very useful ideas for their own institutionally developed surveys.

Readers interested in the COFHE alumni survey should contact: Dr. Larry H. Litten, Consortium on Financing Higher Education, Suite 307, 238 Main Street, Cambridge, MA 02142. Tel: (617) 253-5030. Fax: (617) 258-8280.

Assessment Update
September-October 1992
Volume 4, Number 5

GARY R. PIKE

The University of Missouri Freshman Survey

Dropping out of higher education is extremely costly, both for the student and for the college or university. For the student, these costs include the emotional and intellectual energy invested in obtaining an education, as well as the economic costs associated with attending college. For the institution, the costs of dropping out can be measured in lost income and the expenses associated with recruiting additional students. As John Bean observed, "It takes four freshmen who quit after one year to equal the income of one student who stays for four years" (1990, p. 147).

Traditionally, assessment practitioners and institutional researchers have attempted only after the fact to determine why students leave.

Traditionally, assessment practitioners and institutional researchers have attempted only after the fact to determine why students leave. That is, the traditional method colleges and universities use in collecting data about students who drop out is to wait until the students do not return, then send them a questionnaire about their college experiences and ask why they did not return.

There are at least three potential problems with this approach. The first is that waiting until after students have left the institution may result in a low rate of return for the questionnaire. Research has shown that one of the most important factors in achieving a high rate of return is having accurate mailing addresses. If an institution waits up to three months before sending out a survey, the likelihood of having accurate addresses for nonreturning students declines significantly. Add to this the possibility that nonreturning students have a negative opinion of the college or university and the likelihood of a high response rate drops dramatically. While I was at the University of Tennessee, Knoxville (UTK), we were proud of the fact that we had about a 40% response rate on our survey of nonreturning students.

Another potential problem with waiting to collect data until students have left the institution is that the data are, by definition, retrospective. That is, students are being asked what they did and how they felt three months after the fact. Certainly there is some degradation in the accuracy of the information over three months. Moreover, some of the students who do not return undoubtedly have negative feelings about the institution. There is a very real

Assessment Update
September-October 1995
Volume 7, Number 5

BEST COPY AVAILABLE

possibility that these negative feelings will cause students to rationalize their self-reports on the survey. In effect, a sort of negative halo may permeate the data from the dropout survey.

A third problem with reliance on a retrospective dropout survey is that it provides data about the experiences of nonreturning students only. Comparative data on the experiences of students who return to school in the fall is not available unless a second survey is administered to them.

In 1995 the Office of Student Life Studies at the University of Missouri, Columbia (MU), tried a different approach. During the winter-spring semester, Student Life Studies sent a survey to first-year students enrolled at MU. This survey asked the students to report on their experiences during their first year. The plan is to take the data from the survey and merge it with student records and dropout data from next fall in order to build a data base concerning the characteristics and experiences of returning and nonreturning students. First-year students were surveyed because the highest dropout rate at MU occurs between the first and the second years.

Student identification numbers associated with survey responses made it possible to merge information from student records with the survey responses, thereby providing a more complete picture of the students and their experiences at MU.

The MU survey was divided into four parts. The first part contained a series of questions dealing with students' perceptions of the quality of their experiences at MU overall. Specific items included questions about satisfaction with college, intent to return in the fall, faculty-student interaction, peer interaction, and quality of teaching.

The second part of the survey contained questions about how often students had used various services provided by student affairs departments, such as admissions and records, financial aid, the counseling center, campus dining, and the university bookstore. In addition to indicating how frequently they used these services, students were asked to rate their quality.

The third part of the survey contained a series of questions dealing with students' experiences during their first year at MU. These questions were designed to measure several constructs that have been shown to be important in student retention. For example, questions in the third part of the survey included items dealing with academic integration, social integration, institutional commitment, goal commitment, parental encouragement to remain in school, and perceptions of discrimination.

The final part of the survey asked students to indicate the extent to which their experiences at MU contributed to their learning and development. The items in this part of the survey were intended to measure education outcomes in four areas: communication skills, quantitative skills, critical thinking skills, and personal development.

Students were not asked to provide demographic information on the survey. Instead, respondents provided their student identification numbers. Having student identification numbers associated with survey responses made it possible to merge information from student records with the survey responses, thereby providing a more complete picture of the students and their experiences at MU.

Assessment Update
September-October 1995
Volume 7, Number 5

Because of delays in instrument development and printing, the MU survey was not mailed to first-year students until the last month of the winter-spring semester. In fact, the last follow-up mailing was sent to students just before finals week. Despite the fact that the survey was conducted very late in the academic year, we were able to achieve a response rate in excess of 40%. When the survey is mailed in late February or early March, we expect to have a much higher response rate.

The results of the survey have already started to pay dividends. One of the first sets of analyses performed by Student Life Studies staff was an evaluation of the use and perceived quality of the services provided by student affairs departments. We are in the process of sharing that information with those departments, and thus far the departments have indicated that the information is very helpful to them in assessing their programs.

Student Life Studies is also looking at the relationships between several of the college experience items and questions regarding students' intent to return to MU in the fall. While the results of our analyses generally confirm findings from previous research, the one finding that has surprised me is the strength of some of the indirect effects in our retention model. For example, our research shows that there is a moderately positive relationship between academic and social integration and the intent to persist. On the other hand, the indirect effects of academic and social integration, operating through satisfaction and institutional commitment, are nearly as strong as the direct effects. From my perspective, this finding reinforces what Pascarella and Terenzini (1991) said about the importance of studying the indirect and contingent effects of students' college experiences.

Our research shows that there is a moderately positive relationship between academic and social integration and the intent to persist.

In sum, the approach utilized in the MU survey seems to be a viable method of gathering data to improve retention and student success. Experience to date suggests that this approach yields higher response rates than traditional surveys of nonreturning students, avoids at least some of the problems with halo effects produced by retrospective surveys, and allows assessment professionals to collect comparable data on returning and nonreturning students. Because my next *Assessment Measures* column will be written after we have gathered data on actual persistence, I intend to share some of the results of our analyses with readers. In addition, I will be in a better position to provide an overall evaluation of this survey.

References

Bean, J. P. "Why Students Leave: Insights from Research." In D. Hossler, J. P. Bean, and Associates, *The Strategic Management of College Enrollments*. San Francisco: Jossey-Bass, 1990.

Pascarella, E. T., and Terenzini, P. T. *How College Affects Students: Findings and Insights from Twenty Years of Research*. San Francisco: Jossey-Bass, 1991.

University of Missouri's Survey of Admitted Students

Conversion of admitted students into enrolled students is a key element in enrollment management.

The funnel is a frequently used metaphor for describing the enrollment management process. At the mouth of the funnel is the area (geographical and demographic) that is served by the institution. Of particular interest are the college-bound students within this market and service area because they represent the prospect pool from which most of an institution's applicants, and, eventually, admitted students, will come. At the bottom of the funnel are those students who are retained and graduated from the institution.

In this article, I focus on the top part of the funnel. In particular, I profile one measure that can be used to assess an institution's recruitment and enrollment efforts. The next *Assessment Measures* column will focus on instruments that can be used at the bottom of the funnel to provide data about retention and graduation.

With respect to the top part of the funnel, it is important to understand that the size and shape of the funnel are influenced by the market and service area of the institution, as well as by its mission. For elite public and private institutions, the mouth of the funnel is very broad, but it narrows rapidly because relatively few students are admitted and even fewer actually enroll at these institutions. For regional two- and four-year open-admissions institutions, the mouth of the funnel is not as wide, but its walls are more nearly vertical because larger numbers of prospective students in the service area apply, are admitted, and enroll.

Irrespective of the shape of the enrollment management funnel, conversion of admitted students into enrolled students is a key element in enrollment management. Information about who is enrolling and why, along with information about who is not enrolling and why, can be used to increase or limit class size, evaluate enrollment efforts, and even inform retention efforts.

A popular method of collecting data about prospective students' decisions to enroll or not enroll is to administer a survey to admitted students close

Assessment Update
July-August 1995
Volume 7, Number 4

BEST COPY AVAILABLE

to the time they are making the decision about which college to attend. One of the most often used commercially available surveys is the *Admitted Student Questionnaire* (ASQ), developed by the College Board. Many other institutions have developed questionnaires that represent variations on the ASQ.

In general, surveys of admitted students ask respondents to indicate the importance of certain college characteristics (for example, academic quality, opportunity for cultural and social life, and cost of attendance) in selecting a college. Prospective students also rate the quality of the institution on each of these characteristics. Additional survey questions ask about the importance and nature of people's opinions about an institution, and about the importance and quality of communications from the institution to the prospective students, and from other institutions to which the students have applied. Traditionally, these survey responses have been analyzed by comparing the responses of students who enroll with those of individuals who do not enroll, and results are used to modify the enrollment plan for the coming year.

Last year the University of Missouri-Columbia (MU) used one of these traditional surveys to evaluate its enrollment efforts. While the survey provided a great deal of information to the enrollment management team, it had several drawbacks. First, we felt that certain college characteristics (for example, financial aid) were not given sufficient weight. Likewise, many of MU's enrollment-related efforts (for example, telephone calls from students, faculty, and alumni) were not fully represented in the survey.

The second problem with the survey was its timing. Because the survey was administered late in the college decision process, most of the respondents had already made up their minds. Consequently, the vast majority of the survey respondents had already enrolled at MU. These students reported that everything was important and that MU was the best at everything. This lack of variance in responses made it very difficult to arrive at any firm conclusions about the effectiveness of our enrollment efforts.

The third problem with the survey was the timeliness of the data. Information provided by the survey became available too late to be used to guide that year's enrollment efforts. In the language of evaluation research, we were being forced to use a summative evaluation method when what we really wanted was a formative evaluation that would guide enrollment efforts.

In 1995 we tried something different. Instead of sending out one survey in late April and early May when prospective students are making college choices, we used two surveys, the *Survey of Admitted Students: Part 1* and *Part 2*. The first survey was mailed to students in February and March and included the "importance" items in traditional surveys. The second survey was mailed in early May to prospective students who responded to the first survey. The second survey asked these students to rate MU against other institutions on a variety of college characteristics and enrollment efforts.

Traditionally, survey responses have been analyzed by comparing the responses of students who enroll with those of individuals who do not enroll, and results are used to modify the enrollment plan for the coming year.

The use of a two-part survey has already paid dividends. First, the response rate to the survey was quite good, over 50 percent for students applying before February 1, 1995. (Our response rate for students applying on or after February 1 was extremely low.) The relatively high response rate should ensure that we have the responses of a substantial number of non-enrolled students in the final sample and more variability in responses.

Second, and most important, we have already been able to use the information from *Part 1* to guide our enrollment efforts in 1995. The enrollment management staff called students in certain target groups (for example, high-ability students, particularly minority students) who have been admitted but have not paid their enrollment deposits. Staff members who made these calls were provided with profiles of what students in the target groups reported as important characteristics they looked for in a college. This information allows our callers to target specific messages to prospective students.

One danger of using two surveys is that the response rate to the second is typically lower. For example, a 50 percent response rate to the first survey, coupled with a 50 percent response rate to the second survey, would yield an overall response rate of 25 percent. While this scenario is a very real possibility, we believe that the risk is acceptable given the additional formative information we have gained using this method.

Copies of the *Survey of Admitted Students: Part 1* and *Part 2* can be obtained by writing the Office of Student Life Studies, University of Missouri-Columbia, 3 Parker Hall, Columbia, MO 65211.

T*he use of a two-part survey has already paid dividends.*

Assessment Update
July-August 1995
Volume 7, Number 4

GARY R. PIKE

Issues in Writing Assessment

Since coming to the Center for Educational Assessment (CEA) in January 1993, one of my major projects has been the statewide assessment of student writing for grades K-12. Because many colleges and universities are interested in assessing the quality of student writing as part of their evaluation of general education programs (or their writing-across-the-curriculum programs), I thought that it might be helpful to recount some of our experiences. In this column I will provide a brief description of the writing assessment program in Missouri and discuss some of the basic questions that must be addressed in developing a writing assessment program. In a subsequent column, I will be discussing some of the findings of recent CEA evaluations of the Missouri Writing Assessment program.

Many colleges and universities are interested in assessing the quality of student writing as part of their evaluation of general education programs.

The easiest way to organize this description is to examine three of the key questions that should be answered in order to establish a viable writing assessment program. I want to stress that there are no "right" answers to these questions. However, the ways in which institutions answer them profoundly shape the nature of writing assessment on their campuses.

The first question is "What sort of writing task should be used?" The choice initially is between what I call a one-shot writing assessment and an assessment of the process of writing. In the one-shot assessment, students are given a topic and told that they have a certain amount of time (usually about one hour) to write an essay. This approach is frequently used in commercial general education assessment instruments, such as the Academic Profile and College BASE. The strength of the one-shot method is its ease of administration. Its limitation is that it usually does not reflect the writing process students use.

The alternative, a process-writing approach, is designed to parallel more closely the actual writing process of students. The present Missouri Writing Assessment is an example of the process-writing approach. Students are allotted three class periods in which to write. The first is devoted to prewriting activities and an initial draft. During the second class period,

Assessment Update
September-October 1993
Volume 5, Number 5

students are encouraged to continue drafting and to begin revising their essays. During the final period, students complete their revisions and prepare a final copy of the essay.

Clearly, the major liability of the process-writing approach is the time required for administration and scoring. Scoring takes longer because the essays tend to be longer. However, the time is well spent. Research conducted by CEA has found that the process-writing approach has much greater face validity with classroom teachers than the one-shot approach. In addition, scores tend to be somewhat higher, and there is greater variability in scores. When a one-shot writing assessment was used in Missouri, it was discovered that 85% of the students received a score of 3 on a six-point scale. With the process-writing approach, slightly less than 40% of the students received a score of 3.

Once the nature of the task has been defined, assessment professionals are confronted with the choice of which type of writing prompt to use (for example, personal narrative, expository essay, or persuasive essay). This choice is not trivial. Several researchers have reported that writing scores vary more across types of writing than across students. In 1992, most students in Missouri wrote expository essays. However, a randomly selected subsample of students wrote personal narratives. Results indicated that students who wrote personal narratives had significantly higher scores than students who wrote expository essays (half a point higher on a six-point scale).

Several researchers have reported that writing scores vary more across types of writing than across students.

A related issue in selecting writing prompts is whether the topics should be generic or discipline-specific. Recently, I have been working with several community colleges that are developing assessment programs. One is implementing writing assessment. In pilot testing the writing assessment, it was discovered that standards for writing were significantly different across the disciplines (sciences, social sciences, humanities, and so on). If institutions elect to use a generic writing prompt, they run the risk of not assessing the kinds of writing performance expected of their students. On the other hand, if an institution administers many different kinds of writing prompts, it may be difficult or impossible to generalize across writing samples and arrive at an overall evaluation of the general education writing program.

Irrespective of the type of prompt selected, the directions should include sufficient introductory material to provide the students with a context for the writing exercise. At a minimum, students need to know the purpose for the essay and the intended audience.

The final question that must be answered in developing a writing assessment concerns the type of scoring rubric to be used. For simplicity I will focus on the choice between holistic and analytical rubrics. Holistic rubrics require raters to make a global evaluation of writing samples based on the assumption that the overall impression conveyed by an essay is greater than the sum of its parts. Analytical scoring rubrics contain multiple scoring scales covering different aspects of writing, such as ideas and content,

Assessment Update
September-October 1993
Volume 5, Number 5

organization, word choice, sentence structure, voice, and so forth. Many of the final recommendations suggest that holistic approaches be used for summative end-of-term assessment and that analytical approaches be used for formative assessment designed to provide feedback to students.

Despite the ongoing debate over the relative merits of holistic and analytical scoring methods, one fact is clear. Analytical scoring methods are substantially more time-consuming than holistic ones. For example, one study found that it took 2-1/2 times as long to train raters and score essays using analytical methods. At the present time, CEA and the Missouri Department of Elementary and Secondary Education are conducting a study comparing holistic and analytical scoring approaches. The results of that research will be included in a subsequent column on writing assessment.

It was discovered that standards for writing were significantly different across the disciplines.

Assessment Update
September-October 1993
Volume 5, Number 5

Statewide Writing Assessment (K-12) in Missouri

Missouri introduced statewide direct writing assessment in 1987.

In a previous column (*Assessment Update*, 1993, vol. 5, no. 5, pp. 9, 11), I described several issues that should be addressed when developing a writing assessment program. These include questions about the nature of the writing task (one-shot versus process writing), the prompts (personal narrative, expository, persuasive, and so on), and the type of scoring rubric (holistic, analytical, or a combination) to be used.

In this column, I examine some of the practical issues involved in writing assessment, using as an example the statewide writing assessment at grades 5, 8, and 11 in Missouri. While this illustration is drawn from K-12 assessment, the principles outlined can be transferred to higher education assessment.

Missouri introduced statewide direct writing assessment in 1987. From 1987 to 1989, students wrote for one class period on topics that required a personal narrative style of writing. Approximately 30 teachers were trained as raters and scored the writing samples over the summer using a holistic scoring rubric.

Analyses of the Missouri writing assessment from 1987 to 1989 found that the project had little or no discernible effect on the teaching or learning of writing. Based on these results, an advisory committee met to discuss methods of improving writing and writing assessment in Missouri. The outcome of these deliberations was the development of a process-oriented writing assessment that began in 1990. The purposes of the current writing assessment are threefold: (1) to emphasize the importance of writing, (2) to familiarize students and teachers with the process approach to writing, and (3) to provide data on the quality of writing proficiency in Missouri.

The present writing assessment is organized as a three-day writing workshop that models good instruction and pedagogy in order to familiarize teachers with process-oriented writing strategies and to encourage teachers to develop curriculum-embedded writing assessment. Students are allotted

Assessment Update
January-February 1994
Volume 6, Number 1

three 45-minute class periods to write. The first class period is devoted to prewriting activities and an initial draft. In the second period, students are encouraged to continue drafting and revision activities. During the third class period, students complete the revisions and copy their papers in ink into scannable test booklets.

The Missouri Department of Elementary and Secondary Education (DESE), in conjunction with the Center for Educational Assessment (CEA), coordinates holistic scoring sessions at several sites around the state. Each year a cadre of elementary and secondary school teachers is recruited and trained to score the essays. Two raters read and score each paper independently. A student's score is the average of these two scores. For example, if a student is given a score of 3 by the first rater and of 4 by the second, the final score for that student will be 3.5. When there is more than a 1-point difference in the scores of the two raters, a third rater is asked to read and score the paper. When a third rater is used, the student's score is the average of the two closest scores.

Anchor, training, and certification papers play an important role in the training of raters. These papers are used to familiarize raters with the scoring rubric, provide good examples of each point on the scoring rubric, provide opportunities for raters to employ the rubric prior to the actual scoring sessions, and evaluate whether raters can properly employ the rubric in scoring essays.

The experiences with writing assessment in Missouri should teach us, first, that we must keep the goals of assessment in mind when designing an outcomes assessment program.

Anchor papers are intended to provide the best examples of the score points on the rubric. Initially, DESE and CEA staff meet to select approximately 20 "rough" anchor papers. These are used to train members of the anchor selection committee, which is asked to identify at least 60 papers on which there is strong consensus for inclusion in anchor, training, and certification packets. Papers selected by the committee must have at least 70% agreement on the score for inclusion in certification and training packets. There must be at least 90% agreement on the papers selected for the anchor packets.

Six packets of papers are used to train raters. These packets include two anchor packets representing the score points in the rubric, two practice packets representing most of the score points in the rubric, and two certification packets representing any variety of score points. In addition, the anchor selection committee identifies "red-dot" papers to check the accuracy of raters during the scoring of the Missouri writing assessment.

The training of raters begins with a description of the scoring rubric, and anchor papers are presented as examples of the various score points. Once prospective raters are familiar with the scoring rubric, they are given the practice packets and asked to rate the papers. These practice ratings are discussed and compared to the scores awarded by the anchor selection committee.

Upon completion of the training session, the prospective raters are certified. In order to be certified, a prospective rater must match predetermined scores exactly on at least 50% of the papers and be within one point of the predetermined score on 80% of the papers in the packet. If raters are not certified on the first attempt, they are retrained and recertified. Once certified, raters begin scoring the writing samples.

Red-dot papers are used to maintain the integrity of the rubric during the actual scoring process. A total of four sets of five different red-dot papers is given to raters midway through each of three scoring sessions: the afternoon of the first day, the morning of the second, and the afternoon of the third. Raters score the red-dot papers, and these scores are compared to predetermined scores to evaluate whether decay or drift in scoring has occurred and to identify raters who should be retrained.

When all of the papers have been scored, CEA staff scan the Missouri writing assessment scores electronically, process the data, and return the information to the schools in the form of individual and group summary score reports for specified grade levels. In addition, statewide summaries are prepared for DESE.

Accurate and reliable scoring of writing assessment requires the use of appropriate anchor, training, and certification materials.

This lengthy discussion of K-12 writing assessment in Missouri is warranted not because it is the ideal model for writing assessment but because it illustrates three practical aspects of writing assessment:

Principle 1: There is no substitute for rigor. Too many people seem to feel that subjective performance assessments do not require the same degree of rigor in research methods as do standardized multiple-choice tests. I contend that reliance on subjective evaluations requires greater control because the presence of subjective evaluations adds another source of variance (and possible error) to the assessment process. A clear understanding of the scoring rubric, adequate training, and certification of raters are essential to ensure that score differences reflect differences in students' abilities, not differences in raters.

A concrete example may serve to make this point. In 1993, as part of a special study of the Missouri writing assessment, a subsample of papers from 1992 was rescored using novice raters who were given anchor, training, and certification papers from an earlier prompt. Results indicated that percentages of interrater agreement were much lower in the special study than for the original scoring. In addition, the special-study scores were significantly different from the scores in the original writing assessment. The conclusion from this research was that accurate and reliable scoring of writing assessment requires the use of appropriate anchor, training, and certification materials.

Principle 2: Assessment frequently serves multiple purposes. A corollary of this principle is that multiple purposes produce trade-offs between the goals of the program and methodological rigor. For example, the goals of the Missouri writing assessment include training teachers, improving writing, and obtaining evaluation data. Of these three goals, training teachers in the use and evaluation of process writing is the most important. Excluding

Assessment Update
January-February 1994
Volume 6, Number 1

teachers from the scoring sessions may improve the reliability of the scoring process, but it does not familiarize teachers with process writing. Consequently, certification standards for raters of the writing samples are significantly lower than the certification standards CEA imposes for College BASE, and they are lower than the standards set for other K-12 assessments, such as NAEP. At the same time, research indicates that raters tend to perform to the level of certification (for example, a 50% certification standard translates into 50% interrater agreement). Thus, the decision to set a relatively low certification standard allows virtually all teachers to be certified and to gain experience in evaluating process writing, but it also lowers the reliability of the writing scores.

In any assessment program with multiple goals, certain trade-offs are inevitable. What is important is that assessment leaders explicitly consider these trade-offs as they set priorities and design an assessment program.

Principle 3: Assessment must be coupled with feedback to produce improvements in performance. Early (1987-1989) writing assessments in Missouri used one-shot methods that were not representative of how students write and provided little feedback to students or teachers. It was not surprising that little improvement in writing was found. Subsequent writing assessments have used more authentic methods and provided feedback to students, teachers, and school administrators. For the upcoming writing assessment, both holistic and modified analytical scoring approaches will be used in order to provide students and teachers with additional information about the strengths and weaknesses of students' abilities.

Whether the goal of an assessment program is accountability or improvement, methodological rigor in the conduct of the assessment is essential.

One possible component of the feedback loop is describing and teaching to the rubric. On the surface, this may seem like teaching to the test. However, if the rubric represents good writing practice, then describing and teaching to the rubric serves to make the standards for assessment (and good writing) clear to the students. While this practice may improve writing scores, it also improves the quality of students' writing.

In summary, the experiences with writing assessment in Missouri should teach us, first, that we must keep the goals of assessment in mind when designing an outcomes assessment program. If the primary goal of the assessment is improvement, procedures must be developed to provide accurate and appropriate training and feedback to students, teachers, and administrators. Whether the goal of an assessment program is accountability or improvement, methodological rigor in the conduct of the assessment is essential.

Assessment in the Major at UTK

Both for outcomes assessment and for test development, the process is just as important as the product.

Assessment Update
Summer 1989
Volume 1, Number 2

Despite the proliferation of commercially available tests designed to assess student learning and evaluate education programs, many faculty groups have chosen to develop their own measures of student outcomes, either because tests are not available or because available tests do not match the curriculum of a program. At the University of Tennessee, Knoxville (UTK), more than 45 departmental faculties have developed tests in the major. In this column, I want to offer some suggestions for developing outcomes measures, based on experiences at UTK.

At the outset, I should explain that I view test development as a process of identification, evaluation, and revision. In many ways, test development is a microcosm of the assessment process, in which educational goals are identified, attainment of those goals is measured and evaluated, and educational improvements are implemented. Both for outcomes assessment and for test development, the process is just as important as the product, because the quality of the process helps determine the quality of the product.

Before beginning the test-development process, faculty must decide whether they are going to evaluate students, programs, or both; whether to use recognition (multiple-choice items) or production (essay or performance) measures; and whether to use an absolute (criterion-referenced) or relative (norm-referenced) standard of evaluation. Answers to these questions help guide the test-development process as well as determine the characteristics of the test.

Once the basic questions about the goals of assessment have been answered, an institution may decide to create a committee to oversee test development. The committee may be composed of administrators, faculty members, students, alumni, and employers—representatives of every group likely to be affected by the education program. The oversight committee can make an immediate contribution by helping to develop a list of educational goals and objectives for the program. This list should be written carefully, with consideration given to the following guidelines: the goals and objectives should clearly indicate which knowledge or skill is being covered, should represent important parts of the curriculum, and should indicate what a student must do to demonstrate acceptable performance. The complete list of goals and objectives should be reviewed by the oversight committee and by departmental faculty and should be revised as necessary.

Once a list of goals and objectives has been agreed upon, test developers can begin to create test questions. Much of the literature on test development treats item writing as an art, not a science. Nevertheless, several standardized methods for generating test questions are available, and test developers may wish to examine some of these; a short list of references on item writing is included at the end of this column. More important than how an item is written is the fact that the content of the item should reflect a program's goals and objectives, as well as what is actually taught in the program.

After identification of the goals and objectives of an education program and the development of a set of possible test questions, the process of item evaluation can begin. Evaluation includes both a qualitative and quantitative component. Qualitative evaluation of test items involves matching test questions with goals and objectives, to determine whether the content of the test items matches the content of the objectives. It also may be useful to have faculty members rate the importance of test items and to determine whether the content being measured by the item is covered in the curriculum. Faculty review of test items can even serve as a type of internal program review, focusing attention on the match between program goals or objectives and the curriculum.

Quantitative evaluation of test items should always involve pilot testing the questions and may include large-scale testing as well.

Quantitative evaluation of test items should always involve pilot testing the questions and may include large-scale testing as well. Pilot tests of potential questions can be conducted with fairly small samples of students (50 to 100) and, depending on the number of questions and students, test takers may be asked to respond only to a sample of questions. The goal of pilot testing should be to gather enough data to evaluate the difficulty and discrimination of the items, so that poor items can be rewritten or discarded. Whenever possible, students from several classification levels should be tested, to determine whether the test is sensitive to the effects of instruction. It also may be useful to gather information about the test takers, such as their backgrounds, ability levels, and motivation. All of this information will help developers select the best items for a test.

If possible, developers should conduct large-scale tests of their instruments to evaluate the reliability and validity of the instrument and to gather baseline data about student performance on the exam. Once again, students should be drawn from several educational levels, and additional information about the test takers should be gathered.

Once the evaluation phase is completed and the necessary revisions have been made, the test can be used to assess student outcomes. The process of identification, evaluation, and revision does not stop, however. Student scores are also instrument-evaluation data and should be combined with previously collected information, to identify poor questions and to suggest possible revisions in the test.

Sources on Item Writing

Gronlund, N. *Constructing Achievement Tests* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall, 1982.

Morris, L. L., and Fitz-Gibbon, C. T. *How to Measure Achievement*. Newbury Park, CA: Sage, 1978.

O'Neil, H. F., Jr. *Procedures for Instructional Systems Development*. Orlando, FL: Academic Press, 1979.

Roid, G. H., and Haladyna, T. M. A *Technology for Test-Item Writing*. Orlando, FL: Academic Press, 1982.

Assessment Update
Summer 1989
Volume 1, Number 2

GARY R. PIKE

Project for Area Concentration Achievement Testing

Guest Columnists: Anthony J. Golden, director, and Denis L. Squire, associate director, the Project for Area Concentration Achievement Testing, Austin Peay State University.

The Third Alternative

The utility of nationally standardized examinations for outcomes assessment in the major has been and continues to be debated widely. Typically, the issues revolve around the nature of the group used for comparisons and the degree of concordance between the material contained in a given examination and the material covered in a department's courses. That is, faculty are concerned about whether their departments are measuring student achievement of departmental goals or comparing the goals to some national standard.

Faculty are concerned about whether their departments are measuring student achievement of departmental goals or comparing the goals to some national standard.

In order to address these and other issues, many departments have elected to create locally developed instruments, which are intended to emphasize programmatic strengths and examine potential weaknesses while indicating possible directions for improvement or further investigation. These instruments closely match the department's curriculum, producing results that can be applied directly to the program because there is a direct link between the program, implementation of changes, and effects on students. Some locally developed assessment instruments, however, are so well fitted to the individual department that, although amply descriptive, they are without external validation and relatively insensitive to improvement across student cohorts.

For those departments that have found both nationally standardized and locally developed instruments unsuitable, an alternative has been created by the Austin Peay State University Project for Area Concentration Achievement Testing (PACAT). Begun as a consortium of Tennessee psychology departments in 1984, PACAT became a national project as

Assessment Update
March-April 1991
Volume 3, Number 2

BEST COPY AVAILABLE

a result of funding received in 1988 from the U.S. Department of Education Fund for the Improvement of Postsecondary Education (FIPSE). Responsible initially for producing a cooperative assessment instrument, the Area Concentration Achievement Test (ACAT), the project currently serves 39 academic departments at 28 postsecondary institutions in 12 states. The project's goal is to construct ACATs that conform closely to actual departmental requirements, as determined through national surveys.

During 1988-90, PACAT staff distributed 10,125 surveys to faculty in all 50 states, the District of Columbia, Puerto Rico, and the Virgin Islands, with returns in excess of 4,700. The disciplines included in the surveys were art, biology, communication, criminal justice, literature in English, geology, history, physics, political science, psychology, public administration, social work, and sociology. Surveys currently are being prepared for departments of agriculture, chemistry, general business, and mathematics. The survey results have been used to construct curricular profiles that permit groups of departments with similar content requirements for the major to be matched. To date, approximately 2,000 sets of survey results have been sent to departments requesting them.

Departments can compare their results to a norm containing students from a number of institutions.

On the basis of these results, PACAT works with the faculty of participating departments to construct test items reflecting appropriate content and complexity for graduating seniors. PACAT staff use these items to construct versions of the ACAT that conform to the curricular profiles identified in the surveys. Participating departments may select the profiles they perceive to be most appropriate. Most versions of the ACAT contain a set of common content areas, with two or more additional areas selected by the individual department from a list of available options. At this time, ACATs are available in psychology, social work, political science, and literature in English. New instruments for biology and communication will be introduced in the spring of 1991. Item-writing consortia are also being formed or expanded in art, communication, general business, and history.

The ACAT provides many of the benefits of a nationally standardized examination. As they examine percent-correct scores, departments can compare their results to a norm containing students from a number of institutions. Scores are provided for each of the content areas appearing in the instrument. An overall performance score is also provided. Individual student scores are available upon request. PACAT maintains records of individual departmental performances and provides additional scores based on historical performance within the participating department. In effect, the department's former students serve as the norm group for its present students. Departmental involvement in instrument design is minimal and does not require significant statistical or psychometric expertise. Moreover, participating departments are not required to write extensive banks of items in order to begin using a given instrument.

Assessment Update
March-April 1991
Volume 3, Number 2

In addition to offering some of the benefits of a nationally standardized exam, the ACAT also features many of the characteristics of locally developed instruments. The flexible design of the instrument permits it to parallel a department's curricular emphases. Although individual departments cannot select the specific items that will appear on the ACAT, they are encouraged to review the items and, if they like, to submit new ones to be used in subsequent revisions. Faculty who choose to write items are confronted with a much less formidable task, as a result of the opportunity to use a pool of items contributed by different institutions. The scores based on the performance history of the individual department parallel the information that can be obtained from a similar, locally developed, instrument. Special treatments of the data, including individual response records, can be prepared upon request. PACAT staff also work closely with departments wishing to conduct research using the ACAT or to develop innovative applications of the information produced. A number of research studies have already been conducted.

Further details concerning the program can be obtained by writing to or calling the authors. Survey results and examination copies of the ACATs are available upon written request from PACAT, Austin Peay State University, Box 4568, Clarksville, TN 37044. Tel: (615) 648-7451.

The ACAT also features many of the characteristics of locally developed instruments.

Clearinghouse for Higher Education Assessment Instruments: Survey Results

Only 5% of the respondents indicated that they had no assessment program.

Guest Columnists: Joining the author in writing this column were Jama Bradley and Gwynetta Draper, research associates with the Clearinghouse for Higher Education Assessment Instruments.

With a grant from the Fund for the Improvement of Postsecondary Education (FIPSE), the Clearinghouse for Higher Education Assessment Instruments was established in 1991 at the University of Tennessee, Knoxville. Clearinghouse staff have conducted a national survey to document the current status of assessment in higher education and to provide an indication of the types of assessment techniques currently in use.

In early 1992, clearinghouse staff mailed a questionnaire to the chief academic officers of 3,407 two- and four-year institutions. A follow-up mailing was sent to a representative sample of 500 institutions. To date, responses have been received from 565 institutions, and the profile of institutions in this group is remarkably similar to that of institutions on the original mailing list. Approximately 42% of the respondents represent two-year institutions, 20% represent liberal arts colleges and universities, 27% are from comprehensive universities, 6% represent research universities, and 5% represent other types of postsecondary institutions.

The academic vice presidents and deans were asked to provide information about their efforts in seven areas of assessment: basic skills, general education, major fields, critical thinking, student satisfaction, college environment, and affective development. Two general information questions also were included: (1) "How would you characterize your assessment program?" (2) "How widespread is outcomes assessment on your campus?" Response options for the first question ranged from "no program" to "an established program," and response options for the second question included "campuswide," "not currently campuswide, but expected to be in the future," and "do not plan to have a campuswide program."

Only 5% of the respondents indicated that they had no assessment program, 85% indicated that they were just beginning a program, and 10% said they

Assessment Update
January-February 1993
Volume 5, Number 1

BEST COPY AVAILABLE

had an established assessment program. When asked how widespread outcomes assessment was on campus, only 3% indicated that they did not have campuswide assessment and had no immediate plans to develop such a program.

A majority (77%) of all colleges and universities in the respondent group provided assessment of basic skills using one or more methods. Approximately 57% of the respondents reported use of commercially developed tests, while 33% used locally developed measures. About 7% of the institutional respondents indicated that they use computer-adaptive tests, and 6% reported that they evaluate basic skills with some type of project such as a paper, portfolio, or interview. Four-year colleges and universities reported almost the same amount of testing for basic skills as was reported by two-year institutions. However, the four-year institutions were more likely than the two-year colleges to rely on locally developed measures of basic skills (45% versus 24%, respectively).

Two-year institutions were more likely to provide testing in general education than were four-year institutions. For all types of institutions, general education was most frequently assessed using commercially developed tests: 33% of those responding to the survey reported using commercially developed tests, and 17% said they used a project, portfolio, or interview process to assess their general education programs.

Survey responses indicated that the most neglected area of outcomes assessment is noncognitive development.

Two-year institutional respondents reported less activity in assessing outcomes in the major field than did those at four-year colleges and universities. At 41% of the four-year institutions, outcomes in the major field were assessed; this was the case at just 20% of the two-year colleges. Twenty-four percent of the two- and four-year institutions used projects, portfolios, or interviews to assess programs in the major.

Results of the survey also indicated that faculty at the two-year colleges were less likely to assess critical thinking skills than were faculty at four-year institutions (16% versus 32%, respectively). Currently, assessment instruments for evaluating critical thinking skills are somewhat limited in number and in scope. Development of new instruments may increase assessment activity in this area.

At a majority of the two- and four-year institutions, opinion surveys were used to assess student satisfaction. In addition, at approximately 26% of the two-year and 78% of the four-year institutions, projects, portfolios, or interviews were used to assess student satisfaction. Over 40% of all of the respondents used some type of survey to assess the quality of the college environment.

Survey responses indicated that the most neglected area of outcomes assessment is noncognitive development. Less than 25% of all of the respondents assessed noncognitive outcomes such as leadership skills, understanding of self, understanding of other people and cultures, self-esteem, social responsibility, social values, honesty, and mental and physical health.

Obviously, the importance of assessing these outcomes varied with each institution's mission and its ideal of a well educated student. However, the current focus on teaching, leadership skills, and cultural diversity, along with the emergence of the wellness movement, may produce greater interest in assessing noncognitive development in the future.

At the present time, clearinghouse staff are telephoning survey respondents to obtain more information about the types of instruments used and why those instruments were selected. One finding that has emerged from these follow-up interviews is that, despite calls for multiple outcomes measures, most institutions assessed the various types of outcomes using a single instrument or technique.

Readers interested in additional information about the survey may contact Jama Bradley or Gwynetta Draper, Clearinghouse for Higher Education Assessment Instruments, University of Tennessee, 1819 Andy Holt Avenue, Knoxville, TN 37996-4350. Tel: (615) 974-2350.

Despite calls for multiple outcomes measures, most institutions assessed the various types of outcomes using a single instrument or technique.

Assessment Update
January-February 1993
Volume 5, Number 1

GARY R. PIKE

Clearinghouse for Higher Education Assessment Instruments: Resources

In September 1991, the University of Tennessee, Knoxville (UTK), received a grant from the Fund for the Improvement of Postsecondary Education (FIPSE) to establish the Clearinghouse for Higher Education Assessment Instruments. The original goals of the clearinghouse were to (1) conduct a survey of assessment practice in order to identify the types of information needed about assessment instruments and methods and to evaluate the state of the art in assessment practice; (2) develop a series of recommendations concerning the development and evaluation of assessment instruments and methods; (3) compile and maintain a current listing of commercially and locally developed instruments and methods; (4) develop detailed evaluations of the instruments and methods, including the rationale for development, steps used in development, objectives assessed, sample items, evidence of strengths and limitations, and examples of uses made of the data collected via the instruments and methods; and (5) disseminate the descriptions of the instruments and methods to assessment practitioners.

The materials developed by the clearinghouse can be of tremendous assistance to assessment practitioners.

In 1993, Jama Bradley, Gwynetta Draper, and I presented the results of the survey of assessment practice in this column (see *Assessment Update*, Vol. 5, No. 1). Since that time, clearinghouse staff have been collecting, cataloguing, and reviewing assessment instruments from across the nation. The clearinghouse has produced 12 different catalogues and reviews of assessment instruments and methods. These publications cover assessment of general education, major fields, affective outcomes, portfolios, community college outcomes, and student services and student development.

One of my duties as a member of the advisory panel for the clearinghouse is to evaluate its effectiveness at the end of the grant period (August 1994). I decided to publish a part of my evaluation of clearinghouse operations in this column for two reasons. First, I believe that the materials developed by the clearinghouse can be of tremendous assistance to assessment practitioners, irrespective of whether they are just getting started, are in the process of implementing campuswide assessment programs, or have well-established assessment efforts. Second, I believe that the readers of this

Assessment Update
September-October 1994
Volume 6, Number 5

column can and should make significant contributions to the continuing work of the clearinghouse.

By way of review, I want to enumerate the strengths and limitations of the materials currently available from the clearinghouse. At this point, the greatest strength is the sheer number of assessment instruments its staff have catalogued. For example, one of the most popular clearinghouse publications lists 12 instruments for assessing general education outcomes, 22 basic skills measures, 4 writing assessments, and 21 measures of critical thinking. Even more impressive is the catalogue of measures for assessing student development and student services. Well over 200 instruments are listed in that clearinghouse publication.

In addition to breadth of coverage, the clearinghouse publications are beginning to provide depth of coverage. In the area of general education assessment, for example, instrument descriptions are provided for all of the most frequently used assessment measures. These descriptions have been supplemented with essays written by users of the measures that explain how and why the instruments are used. Also included in some of the essays are descriptions of the positive results associated with instrument use.

How instruments are being used will prove to be the most helpful to assessment practitioners.

While breadth and depth of coverage are the strengths of the clearinghouse reviews, they are also their greatest limitations. I believe that even greater breadth and depth are needed. For example, I expect that the reviews of 12 general education assessment instruments only begin to scratch the surface of the variety of instruments currently being used to assess general education programs. Many of the instruments that have not been catalogued are locally developed measures, and other institutions will not become aware of these instruments unless an organization like the clearinghouse takes the lead in disseminating information about them.

Greater depth in reviews is also needed. While the clearinghouse is accumulating a growing body of literature on how assessment instruments are being used, only superficial information is available about some of the instruments. I believe that the reviews of how instruments are being used will prove to be the most helpful to assessment practitioners. Consequently, I would like to see three, four, or even five descriptions of instrument use included with every measure catalogued by the clearinghouse.

This is where the readers of *Assessment Update* can help. If your institution has developed its own assessment instruments, contact the clearinghouse and provide the staff with information about them. If your institution has not developed new instruments, contact the clearinghouse and provide them with information about how you are using commercially available measures. The larger and more detailed the catalogues that are developed by the clearinghouse, the greater the benefits to the assessment community as a whole.

For more information, please contact Michael K. Smith, Clearinghouse for Higher Education Assessment Instruments, College of Education, 212 Claxton Education Building, University of Tennessee, Knoxville, TN 37996-3748. Tel.: (615) 974-5894, Fax: (615) 974-3748.

Assessment Update
September-October 1994
Volume 6, Number 5

GARY R. PIKE

Enrollment Management

Since I began writing this column in 1989, I have reviewed a variety of instruments that can be used to assess end-of-program outcomes and educational processes. In the last four months, my experience in conducting student research for admissions and financial aid has convinced me that accurate assessment of college inputs using an enrollment management framework can be as important as assessing the processes and outcomes of college. In this column I provide a brief overview of the importance of enrollment management research, and in succeeding columns I will review some of the measures available for this process.

In *The Strategic Management of College Enrollments*, Hossler and Bean (1990, p. 5) defined enrollment management as "an organizational concept and a systematic set of activities designed to enable institutions to exert more influence over their student enrollments." They further noted that enrollment management has two general goals: to control the size of the student body and to exert greater influence on the characteristics of the student body.

Controlling the size of the student body and influencing its characteristics are not ends in themselves. Rather, they are the first steps in a process designed to improve retention, graduation, and student outcomes. Consistent with this view, the chancellor of the University of Missouri, Columbia, stresses the importance of recruiting and enrolling successful graduates.

Bean and Hossler (1990) identified five key elements in enrollment management: (1) using institutional research to position the campus in the marketplace and examine the correlates of student persistence, (2) developing appropriate marketing and pricing strategies, (3) monitoring student interests and academic program demand, (4) matching student demand with curriculum offerings, and (5) identifying factors that can influence retention and graduation. To be successful, enrollment management activities need to be based on empirical data and the findings need to be linked to planning and budgeting processes. It does little good to identify a new

To be successful, enrollment management activities need to be based on empirical data and the findings need to be linked to planning and budgeting processes.

Assessment Update
May-June 1995
Volume 7, Number 3

BEST COPY AVAILABLE

market of potential applicants, or to identify factors related to retention and graduation, if there is no way to implement programs designed to take advantage of enrollment management research results. Moreover, my recent experience suggests that enrollment management research can be most effective when it is linked to decision making at a variety of levels in an institution. Certainly it is important for institution-wide recruitment, conversion, and admissions criteria to be informed by enrollment management research. However, it is also important to provide enrollment management information at the division and department levels, where many of the programs designed to promote student success are actually carried out.

As I noted at the beginning of this column, enrollment management is becoming increasingly important for public and private colleges and universities, and a variety of new research instruments and methods are being made available. In future columns I will be examining some of these instruments and methods and indicating how they can be used to improve the fit between students and the institution. If your institution has developed enrollment management instruments and methods, I would appreciate hearing from you so that this column will be able to provide readers with the state of the art in new enrollment management measures.

Enrollment management is becoming increasingly important for public and private colleges and universities.

References

Bean, J. P., and Hossler, D. "Tailoring Enrollment Management to Institutional Needs: Advice to Campus Leaders." In D. Hossler, J. P. Bean, and Associates, *The Strategic Management of College Enrollments*. San Francisco: Jossey-Bass, 1990.

Hossler, D., and Bean, J. P. "Principles and Objectives." In D. Hossler, J. P. Bean, and Associates, *The Strategic Management of College Enrollments*. San Francisco: Jossey-Bass, 1990.

Assessment Update
May-June 1995
Volume 7, Number 3

To obtain complete back issues of the *Assessment Update* journal, contact Jossey-Bass Publishers for price and availability.

Phone: 1-800-956-7739

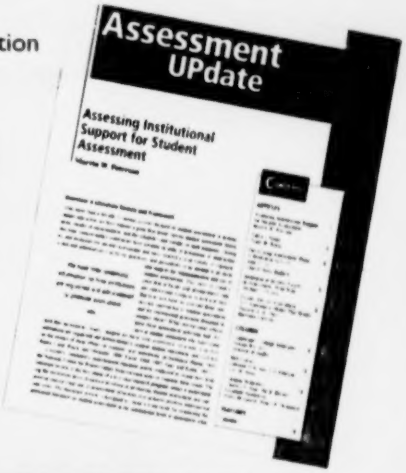
Mail: Jossey-Bass Publishers
350 Sansome Street
San Francisco, CA 94104

Fax: 1-800-605-2665

Website: www.josseybass.com

"In academic organizations dedicated to the continuous improvement of teaching and learning, assessment becomes fundamental and necessary, and *Assessment Update*—a terrific resource for tracking advances in the field—becomes an essential read."

—Theodore J. Marchese, vice president, American Association for Higher Education



ASSESSMENT UPDATE

PROGRESS, TRENDS, AND PRACTICES IN HIGHER EDUCATION

Pressures for assessment are increasing all the time, both from within institutions and from without. In an era of tight money and demands for accountability from a whole range of groups, that's not going to change. But even without imposed mandates, keeping up with the latest developments in assessment is important because assessment is a powerful tool for measuring how well you're meeting the educational needs of your students.

There's only one publication that keeps you at the leading edge of assessment today—***Assessment Update.***

A POWERFUL TOOL FOR INSTITUTIONAL IMPROVEMENT

Assessment Update, a bimonthly newsletter published by Jossey-Bass Inc., has been covering the latest developments in assessment nationwide—even world-wide—longer than any other publication. *Assessment Update* brings it all together in one place, exploring the latest breakthroughs and reassessing the "tried and true." Each 16-page issue features thoughtful articles and columns written by some of the best-known names in assessment and gives you a mix of top-notch scholarship and advice rooted in the practical experience of its authors.

Assessment Update offers you:

- Evaluations of new and established assessment theory and practice
- Advice on conducting assessments in a range of areas, including student learning and outcomes, faculty instruction, general education, academic programs and curricula, student services, and more
- Innovative program models from today's universities and community colleges
- Examinations of the various types of state mandates for assessment

THE EDITORIAL BOARD OF *ASSESSMENT UPDATE*

Editor

Trudy W. Banta, vice-chancellor for planning and institutional improvement, and professor of higher education at Indiana University–Purdue University, Indianapolis

Managing Editor

Karen Elaine Black, assistant to the vice-chancellor at Indiana University–Purdue University, Indianapolis

Assistant Editor

Frances W. Oblander, director of institutional research, planning, and effectiveness, Owensboro Community College

Consulting Editors

Peter T. Ewell, senior associate at the National Center for Higher Education Management Systems

Thomas A. Angelo, associate professor and coordinator of higher education programs, University of Miami

Marcia Mentkowski, professor of psychology and director of the Office of Research and Evaluation Alverno College

Jeffrey A. Seybert, director of research, evaluation, and instructional development, Johnson County Community College

100% Money-Back Guarantee

BEST COPY AVAILABLE

WHAT YOU'LL FIND IN *ASSESSMENT UPDATE*

ARTICLES OF ENDURING VALUE

Each issue of *Assessment Update* presents three to five articles by leading experts and assessment practitioners, exploring the diverse purposes, methods, benefits, and potential pitfalls of assessment, and giving advice for successfully applying assessment results for institutional enhancement.

Recently published articles include:

- Closing the Feedback Loop in Classroom-Based Assessment, by Barbara E. Walvoord, Barbara Bardes, and Janice Denton
- Assessing Institutional Support for Student Assessment, by Marvin W. Peterson
- Managing the Assessment Process: A Software Solution, by Michael C. Choban and Mark DeFoe
- The Performance Funding Wave: Views of State Policymakers and Campus Leaders, by Andreea Serban
- Alternative Assessment: Popularity, Pitfalls, and Potential, by Sheila C. Ewing
- Helping Teachers Define and Develop Authentic Assessment and Evaluation Practices, by Rita A. Moore
- The Cal State Trustees Remediation Policy: Assessment as Part of the Problem and Part of the Solution, by Peter S. Hoff
- Finding Balance in General Education Assessment, by Bill Wresch

COLUMNS BY TOP EXPERTS

Campus Profiles, by Peter J. Gray—Examines a single institution's assessment program.

Assessment Measures, by Gary R. Pike—Evaluates new and existing tests, surveys, performance measures, and other instruments.

From the States, by Peter T. Ewell—Looks at how state and federal mandates are affecting assessment practice.

Web Corner, by Ephraim Schecter, Cel Johnson, and Alec M. Testa—Guides readers to the latest assessment resources available on the Internet.

Community College Strategies—Discusses assessment procedures, problems, and solutions unique to two-year institutions.

Campus Strategies—Focuses on particular assessment methods, reported on by the people who developed and implemented them.

NEWS IN BRIEF

Memos Brief reports on research results, new assessment initiatives, experimental instruments, and other developments.

Resources Reviews of new books, articles, videos, and other resource materials.

Calendar Lists upcoming regional and national conferences and meetings.

YES! I want to subscribe to *Assessment Update*—$75 for six issues.

☐ Payment enclosed (N.Y. residents only, please include sales tax. Canadian residents, add GST and local taxes. For orders outside the U.S., please add $25 for airmail)

☐ VISA/MasterCard/AmEx/Discover

Card No. ______________________ Exp. date ____________

Signature ______________________ Day phone ____________

☐ Bill my organization (**Purchase order and phone number required.**)

P.O. No. ______________________ Day phone ____________

Name ______________________________________

Address ______________________________________

____

City/State/Zip ______________________________________

MONEY-BACK GUARANTEE:
If at any time you are dissatisfied with *Assessment Update*, for any reason, simply notify us in writing and we will refund the entire subscription price. No questions asked.

MAIL TO:

Jossey-Bass Publishers
350 Sansome St.
San Francisco, CA 94104
Call: 888.378.2537
or 415.433.1767
Fax: 800.605.2665

Priority code: AUB

Visit our home page on the World Wide Web: http://www.josseybass.com

BEST COPY AVAILABLE

NATIONAL CENTER FOR HIGHER EDUCATION MANAGEMENT SYSTEMS
P.O. BOX 9752 • BOULDER, COLORADO 80301-9752
WWW.NCHEMS.ORG

BEST COPY AVAILABLE 530